

Volume 22, Number 1 ♦ February 1998

Evaluation Review



*A Journal of Applied
Social Research*

**Special Issue:
ZERO EFFECTS OF
DRUG PREVENTION PROGRAMS:
Issues and Solutions**

Edited by Ita G. G. Kreft and Joel H. Brown



SAGE Periodicals Press

F
2
92z
998
INDESMITH

EVALUATION REVIEW

A JOURNAL OF APPLIED SOCIAL RESEARCH

Editor: Richard A. Berk, *University of California, Los Angeles*

Advisory Editors:

Leonard Bickman, *Vanderbilt University*
Robert Boruch, *University of Pennsylvania*
Ken Land, *Duke University*
John B. McKinlay, *New England Research Institute*

Roberta Miller, *CIESIN*
Richard Rockwell, *University of Michigan*
Peter H. Rossi, *University of Massachusetts*
Ann D. Witte, *Wellesley College*

Associate Editors:

Paul Abramson, *University of California, Los Angeles*
Laurie J. Bassi, *Georgetown University*
Henry Brady, *University of California, Berkeley*
William A. Busching, *Aid Association for Lutherans*
Valerie Caracelli, *U.S. General Accounting Office*
James R. Coldren, Jr., *University of Illinois at Chicago*
Paul V. Ellefson, *University of Minnesota*
L. J. Gorenflo, *Argonne National Laboratory*
Robert Gould, *University of California, Los Angeles*
Karl Hammer, *Vanderbilt Institute for Public Policy Studies*
Lawrence C. Heilman, *Management Systems International*
Eric M. Larson, *U.S. General Accounting*
Jerome S. Legge, Jr., *University of Georgia*
Alisa Lewin, *University of Haifa*
Tony Lin, *Hughes Aircraft*
Sharon K. Long, *Urban Institute*
Katherine Marconi, *Human Resources & Services Administration*
Mary Marvel, *Ohio State University*

Daniel Nagin, *Carnegie Mellon University*
Steven D. Pinkerton, *University of California, Los Angeles*
Hallie Preskill, *University of St. Thomas*
Ian Rawson, *Carnegie Mellon University*
Senta Reizen, *National Center for Improving Science Education*
Jerome Sacks, *National Institute of Statistical Sciences*
Allen L. Schirm, *Mathematica Policy Research*
Richard K. Scotch, *University of Texas, Dallas*
Steven P. Segal, *University of California, Berkeley*
Rod Simpson, *Griffith University, Australia*
Herbert L. Smith, *University of Pennsylvania*
Todd Strauss, *Yale School of Management*
Wm. Thomas Summerfelt, *Vanderbilt Institute for Public Policy Studies*
Dee R. Wernette, *Argonne National Laboratory*
John S. Western, *The University of Queensland*
John R. Wilmoth, *University of California, Berkeley*

Editorial Assistant:

Vanessa Beddo

For Sage Periodicals Press:

Marielle Horton, Eric Law,
Rose Tylak, and Nancy Hale

Community Based Programs
Drug Education

Prevention programs
Program Evaluation

NF
8.2
K92z
1998

EVALUATION REVIEW

A JOURNAL OF APPLIED SOCIAL RESEARCH

Volume 22, Number 1

February 1998

Special Issue: Zero Effects of Drug Prevention Programs:
Issues and Solutions

Edited by: Ita G. G. Kreft and Joel H. Brown

CONTENTS

Introduction to the Special Issue:

Zero Effects of Drug Prevention Programs: Issues and Solutions

ITA G. G. KREFT and JOEL H. BROWN

3

100 Years of "Just Say No" Versus "Just Say Know":

Reevaluating Drug Education Goals for the Coming Century

JEROME BECK

15

An Illustration of Item Homogeneity Scaling and Multilevel Analysis Techniques in the Evaluation of Drug Prevention Programs

ITA G. G. KREFT

46

Zero Effects in Substance Abuse Programs:

Avoiding False Positives and False Negatives in the

Evaluation of Community-Based Programs

DAVID RINDSKOPF and LEONARD SAXE

78

The Other Side of the Story:

*Student Narratives on the California Drug, Alcohol,
and Tobacco Education Programs*

MARIANNE D'EMIDIO-CASTON and JOEL H. BROWN

95

The Irrelevance of Evidence in the Development of School-Based Drug Prevention Policy, 1986-1996

D. M. GORMAN

118

Sage Periodicals Press



A Division of SAGE Publications, Inc.
Thousand Oaks • London • New Delhi

EVALUATION REVIEW is the forum for researchers, planners, and policymakers engaged in the development, implementation, and utilization of studies aimed at the betterment of the human condition. The Editors invite submission of papers reporting the findings of evaluation studies in such fields as child development, health, education, income security, manpower, mental health, criminal justice, and the physical and social environments. In addition, **EVALUATION REVIEW** will contain articles on methodological developments, discussions of the state of the art, and commentaries on issues related to the application of research results. Special features will include periodic review essays, "research briefs" (reports on ongoing or completed studies), and "craft reports" (focused on innovative applications of specific evaluation research techniques, concepts, and ideas).

MANUSCRIPTS submitted for publication in **EVALUATION REVIEW** should be sent in triplicate (one original and two clear copies) to the Editor, Richard A. Berk, **EVALUATION REVIEW**, Department of Sociology, UCLA, Los Angeles, California 90024. All copy, including indented matter, notes, and references, should be typed double-spaced on standard white paper. Each manuscript copy should include all tables and figures (on separate pages) and an abstract of no more than 100 words. Footnotes and an alphabetical list of references should be appended to each manuscript copy and should follow the *Chicago Manual of Style* (14th ed.). Referees will evaluate submitted manuscripts anonymously. Therefore, potential contributors should attach a cover sheet to each manuscript copy giving title, author, affiliation, and full mailing address; the body of the manuscript should bear **only the title** of the paper as a means of identification. A detailed style sheet is available upon request to the Editor of **EVALUATION REVIEW**. A copy of the final revised manuscript saved on an IBM-compatible disk should be included with the final revised hard copy. Submission of a manuscript implies commitment to publish in the journal. Authors submitting manuscripts to the journal should not simultaneously submit them to another journal, nor should they submit manuscripts that have been published elsewhere in substantially similar form or with substantially similar content. Authors in doubt about what constitutes prior publication should consult the editor.

EVALUATION REVIEW (ISSN 0193-841X) is published six times annually—in February, April, June, August, October, and December—by Sage Publications, Inc., 2455 Teller Road, Thousand Oaks, CA 91320; telephone (805) 499-0721; fax/order line (805) 499-0871; e-mail order@sagepub.com; <http://www.sagepub.com>. Copyright © 1998 by Sage Publications, Inc. All rights reserved. No portion of the contents may be reproduced in any form without written permission of the publisher.

Subscriptions: Regular institutional rate \$280.00 per year. Individuals may subscribe at a one-year rate of \$81.00. Add \$12.00 for subscriptions outside the United States. Single-issue price: \$16.00 for individuals; \$49.00 for institutions. Orders from the U.K., Europe, the Middle East, and Africa should be sent to the London address (below). Orders from India and South Asia should be sent to the New Delhi address (below). Noninstitutional orders **must** be paid by personal check, VISA, or MasterCard.

Periodical postage paid at Thousand Oaks, California.

This journal is abstracted or indexed in **Automatic Subject Citation Alert**, **caredata ABSTRACTS**, **caredata CD**, **Criminal Justice Abstracts**, **Current Contents**, **Current Index to Statistics**, **Health Instrument File**, **Human Resources Abstracts**, **NCJRS (National Criminal Justice Reference Service)**, **Psychological Abstracts**, **Public Affairs Information Service**, **Sage Public Administration Abstracts**, **Social Planning/Policy & Development Abstracts**, **Social Sciences Citation Index**, **Social Sciences Index**, **Social Work Abstracts**, **Sociological Abstracts**, **Sociology of Education Abstracts**, **SRM Database of Social Research Methodology**, **Up-To-Date Library**, and **Urban Affairs Abstracts**, and is available on microfilm from University Microfilms, Ann Arbor, Michigan.

Back Issues: Information about availability and prices of back issues may be obtained from the publisher's order department (address below). Single-issue orders for 5 or more copies will receive a special adoption discount. Contact the order department for details. Write to the London office for sterling prices.

Inquiries: Address all correspondence and permissions requests to SAGE PUBLICATIONS, Inc., 2455 Teller Road, Thousand Oaks, California 91320. Inquiries and subscriptions with ship-to addresses in the U.K., Europe, the Middle East, and Africa should be sent to SAGE PUBLICATIONS Ltd, 6 Bonhill Street, London EC2A 4PU, United Kingdom. Ship-to addresses in India and South Asia, write to SAGE PUBLICATIONS INDIA Pvt. Ltd, P.O. Box 4215, New Delhi 110 048 India. Other orders should be sent to the Thousand Oaks office.

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Sage Publications, Inc. for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the base fee of 50¢ per copy, plus 10¢ per copy page, is paid directly to CCC, 21 Congress St., Salem, MA 01970. 0193-841X/98 \$.50 + .10.

Advertising: Current rates and specifications may be obtained by writing to the Advertising Manager at the Thousand Oaks office (address above).

Claims: Claims for undelivered copies must be made no later than twelve months following month of publication. The publisher will supply missing copies when losses have been sustained in transit and when the reserve stock will permit.

Change of Address: Six weeks' advance notice must be given when notifying of change of address. Please send old address label along with the new address to ensure proper identification. Please specify name of journal. POSTMASTER: Send address changes to: *Evaluation Review*, c/o 2455 Teller Road, Thousand Oaks, CA 91320.

INTRODUCTION TO THE SPECIAL ISSUE

Zero Effects of Drug Prevention Programs: Issues and Solutions

JOEL H. BROWN

Center for Educational Research and Development

ITA G. G. KREFT

California State University, Los Angeles

INTRODUCTION

Since the middle 1980s, when the U.S. drug war switched into high gear, youth drug prevention or prevention education research has been conducted under a fearful miasma. By attempting to engage in critical scientific discourse, many researchers have been labeled as "soft on drugs" or "drug legalizers." Without engaging on that level in this special issue, they have experienced censure and professional damage that is documented in a recent magazine article (Glass 1997). The censoring aspect of this miasma has left us with a paucity of important information that might otherwise facilitate scientific discourse. And in the interim, the well-being of youths has become all but a rhetorical tool, where the program evaluators seem to forget about the actual experience of the students in these programs, experiences that seem to have been detrimental in some instances (see D'Emidio-Caston and Brown 1998 [this issue]). It has taken until now to simply appreciate the gravity of the youth experience in these programs and to initiate a critical scientific discourse. The research in this issue represents one of the first coordinated attempts to challenge the myopic approach to program evaluation (explicated in this special issue) and to look at the broader issues of students' well-being. All articles are written by independent researchers, who are not evaluators of their own new programs and, for the most part, outsiders of this field.

While editing the articles presented in this issue, several topics emerged. Each topic is of direct relevance to drug prevention and is related to how we conduct and report drug prevention research. The topics discussed in this introduction are based on general and pervasive aspects of this research that

the authors in this issue found in reports of drug prevention programs and in reports by the popular press. The topics are

- Myopic evaluation approaches
- Selective reporting of findings instead of other ones
- Questionable results being given unwarranted status in the popular media
- Masking detrimental program effects
- Performing upside-down science: Assuming effectiveness rather than testing for it

MYOPIC EVALUATION APPROACHES

The United States General Accounting Office (GAO) estimates that the federal government is spending about \$2.4 billion annually on youth drug prevention programs (U.S. General Accounting Office 1997). Because of additional state, local, and charitable contributions, by the GAO's own estimates, this figure substantially underestimates the total dollars actually spent on programs. Despite massive spending, though, one glaring fact stares us in the face:

Recently, adolescent substance use has increased more quickly to higher levels than at any time in the past 15 years (Johnston, O' Malley, & Bachman, 1995). Usage increases occur among those youth who have received more drug education than any group since school-based drug education began. (Brown, D'Emidio-Caston, and Pollard 1997, 65)

It seems to be that the current programs operate on the belief that we have the knowledge to successfully prevent youths from using drugs, by transferring the message "Just Say No" into the heads of youths. How to implement this message successfully is seemingly the only problem that needs to be solved. The evidence presented here, as well as in a growing body of literature cited throughout the articles in this issue, however, reveals that current programs and their conceptually flawed underpinnings cannot consistently prevent youths from using or abusing substances.

During the past 20 years, the vast evaluation literature on youth drug prevention appears highly differentiated, yet it is really quite uniform. Research has found that, despite the increase in apparent programmatic complexity, drug prevention programs and their evaluations have changed little (Brown and Horowitz 1993). Consistent with past and current federal program mandates,

nearly all programs implemented in the United States are variations on what the GAO refers to as "no-use" prevention programs (U.S. General Accounting Office 1991). Educators delivering federally funded prevention programs, which means most programs, must teach youths that alcohol, tobacco, and other drugs are harmful and use of any of these substances equals abuse of these substances. Based on these "no-use" or "zero tolerance" mandates, programmers attempt to achieve these goals in three ways (Brown, D'Emidio-Caston, and Pollard 1997). First, providing a no-substance-use message that appeals to youth fears, that is, this picture of a cancerous lung is what will happen to you if you smoke cigarettes. Second, offering rewards (e.g., Drug Abuse Resistance Education [D.A.R.E.] T-shirts) in exchange for promises not to use substances. Finally, attempting to improve students' self-esteem by teaching them about life skills or how to refuse substances when these are offered in a variety of circumstances.

Program evaluations parallel the same overly simplistic ideas found in zero-tolerance programs as described above. The primary outcome measures are simply a delayed onset of, or decreases in, substance use, where substance use is not differentiated between more or less harmful dosage or more or less addictive drugs.

For many years, program success has been determined primarily by the extent to which one point of view (that all drugs are harmful) is affirmed, and youths act accordingly (say no to substance use). The bottom evaluation line is whether youths have sufficiently negative knowledge, attitudes, beliefs, and behaviors about drugs, as a result of prevention programs. The operational definition of program success is exemplified in the survey items, where students, in one form or another, are often asked: "How harmful do you think this drug or that drug is?" Even with these simplistic program concepts and goals, the evaluations of the programs have found very few positive effects (Schaps et al. 1981; Tobler 1986, 1992; Klitzner 1987; Bangert-Drowns 1988; Bruvold 1990; Brown and Horowitz 1993; Ennett et al. 1994; Clayton, Catarello, and Johnstone 1996; Brown, D'Emidio-Caston, and Pollard 1997).

From an educational psychological perspective, we observe that traditional evaluations operate from a narrow definition of influence and learning, regardless of the type of substance or the user's context. There is a single message, and only one type of learning, where the message is "just say no," and the program teaches *what* decision to make, instead of *how* to make a reasoned decision. Such a single-minded message rarely produces long-lasting results, for example, youth abstinence (Raven 1965, 1993; Brown, D'Emidio-Caston, and Pollard 1997). From these perspectives, then a significant shortcoming of

current evaluation methods is that in relation to program implementation, they largely fail to assess the influence on and the learning of the participants. Based on the research of D'Emidio-Caston and Brown (1997), we find that program failure or success needs a wider definition. Students interviewed in this study make a strong case for types of program evaluations in which different substances, licit and illicit, are treated differently, based on the type of substance and level of use, relative to different social contexts. Youths as well as many scientists view these discussions as part of a broader youth development approach.

FAVORING AND REPORTING CERTAIN FINDINGS INSTEAD OF OTHER ONES

In addition to myopic evaluation approaches, there is another concern, which may be a contributor to program failure. There exists a tremendous pressure to favor program success and report positive results. An example was recently found in a 1995 *Journal of the American Medical Association* (JAMA). In what some refer to as the cutting edge of prevention, social influence programs, here termed "life-skills" programs, in a 6-year follow-up of that program, researchers concluded the following:

Drug abuse prevention programs conducted during junior high school can produce meaningful and durable reductions in tobacco, alcohol, and marijuana use if they (1) teach a combination of social resistance skills and general life skills, (2) are properly implemented, and (3) include at least 2 years of booster sessions. (Botvin et al. 1995, 1106)

A close examination of this publication, and after some manipulation with the reported data, shows that alcohol use increased after implementation of these programs in one out of two conditions (the teacher training condition). Amazingly, the negative program effect was not mentioned in the result section, only the positive effect of the other condition. This makes the strong argument for long-term success, as cited above, limited. The article by Kreft in this issue shows similar contradictory findings over analyses of the same data by different people, whereas Brown and Horowitz (1993) demonstrated that, in a more general sense, there exists a clear and consistent pattern of biased reporting, biased toward favorable program effects. It is clear that this pattern has implications for understanding program effects as well as suggesting future program directions.

QUESTIONABLY POSITIVE RESULTS ARE GIVEN UNWARRANTED STATUS IN THE POPULAR MEDIA

In today's information age, the few positive results that are found may be escalated to an unwarranted status. For example, a recent *Time* article told us that the nation's most implemented program, D.A.R.E., was ineffective. At the same time, it extolled the virtues of a new program, the Life Skills program, by telling us to "Just Say Life Skills" (Van Biema 1996, 70). Few know that, absent the police uniform, the methods used to educate kids about substance use in life skills are similar to D.A.R.E., a program that has indeed repeatedly been shown ineffective in the long term (Ennett et al. 1994; Clayton, Catarello, and Johnstone, 1996). From comprehensive reviews of social skills programs, in which life skills is a part, it is learned that, "The majority of studies show that social skills training programs, while not detrimental, have little or no impact upon participants in terms of their alcohol use behavior" (Gorman 1996, 191; see also Gorman 1998 [this issue]).

The news item in *Time* reflects yet another pattern, a pattern observed many times when it comes to a social problem like youth substance use or abuse. The article creates a public uncertainty and later proceeds to fill that uncertainty with reassurance (Skager and Brown forthcoming). This pattern of creating simultaneous public uncertainty (D.A.R.E. may not work) and reassurance ("Just Say Life Skills") tells the public that one knows how to resolve the problem with an apparently new program. But new evidence is telling us that this approach may be a dangerous one; that in addition to poor evaluation methods, biased reporting, and questionable public exaltation, we are masking other detrimental youth program effects.

MASKING DETRIMENTAL YOUTH EFFECTS

As has occurred nearly every time during the past 20 years that youth substance use has risen, researchers have drawn this conclusion: the rises in substance use were due to the rise in youth perception that drugs weren't as harmful for them as they had previously believed (Johnston, O'Malley, and Bachman 1995). We have two examples where this easy explanation is discarded, and where systematic issues associated with drug prevention programs are examined. The first example looks at "at-risk" youths, who are served least by current policies. Previously, programs targeted toward at-risk youths had been viewed as a necessary part of successful prevention programs (Hawkins et al. 1987). Predicting youth substance use is tenuous (see also

Kreft in this issue) but offers the possibility of clear policy and practice directions, which makes the search to connect risk factors with substance use and/or abuse rewarding for many researchers. Yet, during the past 20 years, this research has produced few if any predictive results. Moreover, by linking policy, perception, and practice in a new evaluation research design, researchers found that merely viewing most youths as being at risk for substance abuse caused programmers to enact programs in which those most in need of assistance were often those first removed from the school system (Brown and D'Emidio-Caston 1995). Interview data suggested that student removal occurred through detention, suspension, and expulsion. Despite massive prevention programming efforts targeted specifically toward at-risk youths before being removed, they received programs identical to their thriving counterparts. These authors have stated that this intensive research direction masked how programs were adversely affecting the lives of precisely those youths the programs intended to help. A less myopic and wider approach to program evaluation reveals a group of students that are harmed by these programs.

The second example shows the cognitive dissonance that arises in the mind of students results from the no-use message delivered by the programs. These unintended effects on some youth can be detrimental to drug prevention programs' success. Most current models teach youth how to make the "right" decision, which is "not to use any substances." As a result of the direct conflict between "no-use, all drugs are harmful" messages youth receive in prevention programs verses the multiple levels of use and effects they see outside of school (Brown, D'Emidio-Caston, and Pollard 1997), it has been found that cognitive dissonance with respect to this message arises. Dissonance is indicated by tension, anger, and students' perception that adults were lying to them. The study by Brown et al. reports evidence that their highly negative program perceptions hid far more than mere adolescent rebellion. Youth made clear linkages and logically coherent statements between programs and their feelings about substances and substance use. And educators, often ill trained to deal with these emotional issues, leave many youngsters in this dissonant state. This significant, serious, and unnecessary psychological tension results in reduced adult credibility.

UPSIDE-DOWN SCIENCE: ASSUMING EFFECTIVENESS RATHER THAN TESTING IT

Of course, there could be many explanations why these failed programs continue unabated. But what could explain such strong researcher assertions

for program success in light of contrary information? In Thomas Kuhn's *The Structure of Scientific Revolutions* (1962), we have found a theoretical explanation. As part of what he called "professionalization," individuals in a particular field come to depend on "further articulation and specification" (p. 23) of a model and on developing "esoteric vocabulary and skills" (p. 64), which may help preserve the model's dominance.

If we apply this Kuhnian view to the field of prevention, we observe that although these programs look different, they are not. Researchers are just busy further articulating a commonly accepted prevention vocabulary and skills, such as embodied in the life skills example. On first glance, this approach might appear logical and easy to explain to the public who perceives a deep problem and exerts pressure for an immediate solution. On second glance, though, one result is that in its current state, rather than objectively determining program effectiveness, many evaluators seem to make an extra effort to provide evidence of effectiveness (Brown, D'Emido-Caston, and Horowitz 1996). The result is that nearly all no-use programs are pre-believed to be a success unless proven beyond nearly any doubt that they are not. According to Kuhn, this level of "professionalization leads to an immense restriction of the scientist's vision and to a considerable resistance to paradigm change" (p. 64). This is a phenomenon often observed in prevention. Based on the development of various forms of a no-use model, and favoring certain results over others, we clearly see a resistance to real change.

Whereas Kuhn might call this a resistance to paradigm change, we call it "upside-down" science, or in quantitative terms "the defense of the alternative hypothesis." In traditional quantitative research, the statistical test does not involve the alternative hypothesis. The test protects the null hypothesis. If a comparison of a program with a control group shows not enough evidence to reject the null hypothesis (the supposition that the program effects are zero), the null hypothesis is retained. In prevention, though, many seem to accept the premise that if the test does not show results, the alternative hypothesis needs to be defended. Various explanations are given why this particular intervention did not show results but is nevertheless potentially successful. The hunt for significance is exemplified by many significance tests with different variables and reporting that "one out of 20" chance to find a significant effect, even if the null hypothesis is true.

Moskowitz (1993, 1) summarized the relationship between upside-down science and external researcher pressures quite well when concluding that many "outcome evaluations do not stand up to scrutiny" because of "institutional pressures involved in conducting 'soft-money' research as well as academic pressures to publish or perish and conflict of interest."

OVERVIEW OF THE ARTICLES IN THIS ISSUE

In this section, we discuss the findings in two quantitative and three qualitative examinations of drug prevention programs. Four manuscripts present recent evidence regarding school- or community-based programs. One presents new historical evidence regarding previously unexplored socio-historical developments since the late 1800s. After providing an overview of the findings, we conclude the introduction with a brief discussion of potential programmatic and evaluation solutions in these articles.

In the article by Beck, a social history of drug prevention efforts is discussed. Contrary to the popular belief that today's prevention education was directly rooted in the late 1960s, the historical record suggests that programs strikingly similar to today's programs were initiated more than 100 years ago. Although current programs may appear to be different than those of the past, he shows that in fact, past and present are bound together by more than 100 years of mandatory no-use dictates. His evidence also suggests that current program criticisms can be linked with criticisms that emerged at the turn of the 20th century.

The article by Kreft reexamines a published large-scale evaluation that includes results from one of the many U.S. prevention programming trends, referred to as "normative programs" (Hansen and Graham 1991). The article shows that different ways of analyzing the data may lead to different conclusions. Using the state-of-the-art data analysis techniques, Kreft shows that the traditional way of aggregating data to the level of the class hides more than it reveals. She illustrates what has been well known since Robinson's article appeared in 1950. Analyses executed at an aggregated class level can yield different and even opposite results from data analyzed at the individual level. In the same article, Kreft shows that using dichotomized response variables (use versus non-use) can also lead to very different interpretations of program effects compared to a response variable that shows a more elaborated scale of substance use. The differences in data handling and analyses procedures result in the conclusion that previously drawn conclusions significantly overestimated program effects. Kreft also confirms findings that environmental factors such as adult alcohol abuse in youth's direct environment can predict youth alcohol use.

Rindskopf and Saxe examine a large community-based program. The researchers discuss how they attempted to minimize two types of potential analytical errors: (a) deciding a truly ineffective program is effective (false positive, corresponding to a Type I error), and (b) deciding a truly effective program is ineffective (false negative, corresponding to a Type II error). To minimize the probability of each type of error, several design and statistical

analysis strategies are presented in this community-based evaluation. Although the final results are not in, even in these cutting-edge community programs, when minimizing the error probabilities, the researchers report finding no program effects.

The voices of youths are found in the D'Emidio-Caston and Brown article. They found that narrative and story can be a valuable evaluation tool. Here, the logic of the respondents was found to be consistent with youth decision-making literature showing that many are capable of decision making on par with adults. Students discussed many reasons why drug prevention messages failed, among them the perception that people delivering drug prevention programs do not make a distinction between substance use and abuse, whereas, based on their own observation and daily experiences, most youth do. Moreover, it was found that strict punitive policies and practices alienated those most in need of help. These significant differences between school programs and real-life experiences resulted in their dismissal of programs as not being credible sources of information or assistance. As previously discussed, typical no-use approaches also resulted in youth cognitive dissonance.

From a policy perspective, the Gorman article focuses on the development of school-based drug prevention at the height of the drug war, from 1986 to 1996. In using expenditure and national survey data, he finds little justification for the massive infusion of money into school-based programs. This is achieved by comparing expenditure and survey data before and during the study period. He also shows not only that the expenditures were unnecessary but also that the selected programs were (are) largely unworthy of massive funding and dissemination across the country.

IMPROVING PROGRAMS AND THEIR EVALUATIONS

What do the articles in this special issue tell us about what can be done to improve programs? Much of it has to do with our own attitude toward prevention and, in turn, youths. If, for example, the "Just Say No" message is not a realistic one, if no-use methods do not work, nor do their substrates that teach so-called social and life skills, we have to investigate what needs to be changed. The suggestions made here are mainly based on experience with recipients of the program and a few publications (Brown and Horowitz 1993; Brown 1996) and supported in large-scale research. In these studies, it is found that a youth development approach is more successful than an authoritarian approach, by protecting adolescents from harm, including sub-

stance abuse (Resnick et al. 1997; Tierney, Baldwin-Grossman, and Resch 1995).

Our suggestion is to develop programs with the goal to educate young people to be adults in an informed and free democracy, instead of followers of authority. If the message is based on and integrated with what happens in real life, where some students are confronted with drugs on a daily basis, the message should be changed from punitive to helpful. For instance, the message should be, If you have a substance abuse problem (as about 10% of the American population has), you can do something about it. You can always stop (here is how you can do that) or you can always find someone to help you (and here is where these people are). And in the case of prevention, the message should be honest, providing complete information representing a youth development as opposed to a youth deficit prevention approach.

Suggestions of such solutions as made in the Beck article extend the D'Emidio-Caston and Brown article by proposing a pragmatic alternative solution to today's failing efforts, minimizing the potential harm that might arise from the misuse of substances. It is important to note that whereas many researchers evaluate these factors, using legalistic definitions of no-use mandates and program goals, Gorman and Beck in this issue evaluate substance use in the context of establishing social psychological and health-oriented distinctions between substance use and abuse. Gorman and Beck tell us that we need to take a wider program perspective than appears today. Before the next massive round of dissemination, we must be careful to check and see if the programs are worth the investment in and of themselves, as well as their widespread dissemination. One way to do this is to make data available for secondary analyses by researchers not connected to any of the programs.

POSSIBLE SOLUTIONS

Youth desire for a sense of inclusion and thus connectedness with adults as part of family or in school is necessary to any successful endeavor in this area. And this assertion has been born out in a large-scale *Journal of the American Medical Association* publication concluding that "parent-family connectedness and perceived school connectedness were protective against every health risk behavior measure except pregnancy" (Resnick et al. 1997, 823). Because the emphasis is on the affective connectedness rather than imparting social skills, this method of reducing youth harm by developing

their resilience represents a fundamentally different orientation than current prevention programs.

REFERENCES

- Bangert-Drowns, R. L. 1988. The effects of school-based substance abuse education—A meta-analysis. *Journal of Drug Education* 18:243-65.
- Botvin, G. J., E. Baker, L. Dusenberry, E. M. Botvin, and T. Diaz. 1995. Long-term follow-up results of a randomized drug abuse prevention trial in a White middle class population. *Journal of the American Medical Association* 273 (11): 1106-12.
- Brown, J. H., ed. 1996. *Advances in confluent education, Vol. 1: Integrating consciousness for human change*. Greenwich, CT: JAI.
- Brown, J. H., and M. D'Emidio-Caston. 1995. On becoming at risk through drug education: How symbolic policies and their practices affect students. *Evaluation Review* 19 (4): 451-92.
- Brown, J. H., M. D'Emidio-Caston, and J. Horowitz. 1996. Drug education: What works, what doesn't. Panel Members: Presented at a Congressional Forum, sponsored by the Drug Policy Foundation, May, Washington, DC.
- Brown, J. H., M. D'Emidio-Caston, and J. Pollard. 1997. Students and substances: Social power in drug education. *Educational Evaluation and Policy Analysis (EEPA)* 19 (1): 65-82.
- Brown, J. H., and J. E. Horowitz. 1993. Deviance & deviants: Why adolescent substance use prevention programs do not work. *Evaluation Review* 17 (5): 529-55.
- Bruvold, W. H. 1990. A meta analysis of the California school-based risk reduction program. *Journal of Drug Education* 20 (2): 139-52.
- Clayton, R. R., A. M. Catarello, and B. M. Johnstone. 1996. The effectiveness of drug abuse resistance education (Project D.A.R.E.): 5-year follow up results. *Journal of Preventive Medicine* 25 (3): 1-12.
- D'Emidio-Caston, M., and J. Brown. 1998. The other side of the story: Student narratives on the California Drug, Alcohol, and Tobacco Education Programs. *Evaluation Review* 22 (1): 93-115.
- Dukes, R. L., Ullman, J. B., and Stein, J. A. (1996). Three year follow-up of drug abuse resistance education (D.A.R.E.). *Evaluation Review*, 20 (1): 49-66.
- Edelman, M. (1964). *The symbolic uses of politics*. Chicago: University of Illinois Press.
- Ennett, S. T., N. S. Tobler, C. L. Ringwalt, and R. L. Flewelling. 1994. How effective is drug abuse resistance education? A meta-analysis of Project D.A.R.E. outcome evaluations. *American Journal of Public Health* 84 (9): 1394-1401.
- Glass, S. 1997. Don't you D.A.R.E: America's top drug prevention program doesn't work. But you'd better not say so. *The New Republic*, March 3, 18.
- Gorman, D. M. 1996. Do school-based social skills training programs prevent alcohol use among young people? *Addiction Research* 4 (2): 191-210.
- . 1998. The irrelevance of evidence in the development of school-based drug prevention policy, 1986-1996. *Evaluation Review* 22 (1): 116-144.
- Hansen, W. B., and J. W. Graham. 1991. Preventing alcohol, marijuana, and cigarette use among adolescents: One-year results of the Adolescent Alcohol Prevention Trial. *Preventive Medicine* 20:414-30.

- Hawkins, J. D., D. M. Lishner, J. M. Jenson, and R. F. Catalano. 1987. Delinquents and drugs: What the evidence suggests about prevention and treatment programming. In *Youth at high risk for substance abuse*, edited by B. S. Brown and A. R. Mills, 81-131. (DHHS Publication no. ADM 87-1537; reprinted 1990 as ADM 90-1537). Washington, DC: U.S. Government Printing Office.
- Improving America's Schools Act of 1994. Public Law 103-382. U.S. Government Printing Office.
- Johnston, L. D., P. M. O'Malley, and J. G. Bachman. 1995. *National survey results on drug use from the monitoring the future study, 1975-1994*. Rockville, MD: U.S. Department of Health and Human Services, National Institute on Drug Abuse.
- Klitzner, M. D. 1987. *Part 2: An assessment of the research on school-based prevention programs* (Report to Congress and the White House on the nature and effectiveness of federal, state, and local drug prevention/education programs). Washington, DC: U.S. Government Printing Office.
- Kuhn, T. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Moscowitz, J. M. 1993. Why reports of outcome evaluations are often biased or uninterpretable: Examples from evaluations of drug abuse prevention programs. *Evaluation and Program Planning* 16:1-9.
- Raven, B. H. 1965. Social influence and power. In *Current studies in social psychology*, edited by I. D. Steiner and M. Fishbein, 371-82. New York: Holt, Rinehart, Winston.
- . 1993. The bases of power: Origins and recent developments. *Journal of Social Issues* 49 (4): 227-51.
- Resnick, M. D., P. S. Bearman, R. W. Blum, K. E. Bauman, K. M. Harris, J. Jones, J. Tabor, T. Beunring, R. E. Sieving, M. Shew, M. Ireland, L. H. Bearinger, and J. R. Udry. 1997. Protecting adolescents from harm. *Journal of the American Medical Association* 278 (10): 823-32.
- Robinson, W. S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review* 15:351-7.
- Schaps, E., R. DiBartolo, J. Moskowitz, C. S. Palley, and S. Churgin. 1981. A review of 127 drug abuse prevention program evaluations. *Journal of Drug Issues* 11:17-43.
- Skager, R., and J. H. Brown. Forthcoming. Toward a reformation of drug education. In *Drug policy: Psychological, philosophical and legal issues*, edited by J. Fish.
- Tierney, J. P., J. Baldwin-Grossman, and N. L. Resch. 1995. *Making a difference: An impact study of big brothers/big sisters*. Philadelphia, PA: Public/Private Ventures.
- Tobler, N. S. 1986. Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. *Journal of Drug Issues* 16:537-67.
- . 1992. Drug prevention programs can work: Research findings. *Journal of Addictive Diseases* 11 (3): 1-26.
- U.S. General Accounting Office. 1991. *Drug abuse prevention: Federal efforts to identify exemplary programs need stronger design*. Washington, DC: Author.
- . 1997. *Substance abuse and violence prevention: Multiple programs raise questions of efficiency and effectiveness* (Testimony provided on June 24). GAO/T HEHS-97-166. Washington, DC: Author.
- Van Biema, D. 1996. Just say life skills: A new antidrug program outstrips D.A.R.E. *Time*, November 11, 70.

Through comparative socio-historical analysis of American school-based drug education, this review critically examines past perspectives and practices and how they shaped current programs. Among the key findings emerging from this analysis: Contrary to the popular belief that drug education began in the 1960s, its roots actually go back at least 115 years to the advent of compulsory temperance instruction. Although the particular substances targeted by such approaches have changed, the underlying approaches and dominant "no-substance-use" injunction has not. Despite the existence of "informed choice" approaches, throughout much of this period, evaluation efforts continue to be constrained by the limited dictates of "no-use" perspectives. A pragmatic alternative to contemporary "Just Say No" education is offered that strives to minimize potential harm resulting from the uninformed misuse of licit and illicit substances. A unique evaluative strategy designed to assess the effectiveness of this form of "informed choice" or "harm reduction" drug education is discussed.

100 YEARS OF "JUST SAY NO" VERSUS "JUST SAY KNOW"

Reevaluating Drug Education Goals for the Coming Century

JEROME BECK

Center for Educational Research and Development

INTRODUCTION

Typically, the advent of school-based drug education has been situated amid the 1960s drug crisis. Although formal evaluation of such efforts did commence at this time, drug education itself actually dates much further back to the late 19th century. Even within the substance abuse prevention field, there is little awareness that by 1901, every state and territory had passed

AUTHOR'S NOTE: *The author wishes to acknowledge the many constructive comments provided by Marianne Apostolides and three reviewers of earlier drafts of this article. In addition, thanks go to Mr. and Mrs. Beatty of the Frances E. Willard Memorial Library in Evanston, Illinois for their gracious hospitality in assisting the author through the voluminous Woman's Christian Temperance Union literature found there.*

EVALUATION REVIEW, Vol. 22 No. 1, February 1998 15-45

© 1998 Sage Publications, Inc.

legislation mandating some form of "temperance instruction" to be taught in the public schools.

Up until now, the profoundly ahistorical nature of the contemporary anti-drug campaign has inspired little effort to consult the collective wisdom garnered from past experience. This article attempts a partial redress of this shortsightedness by carefully reviewing our long tradition of school-based drug education. In particular, what instructive lessons can be gleaned from our experience? How have various perspectives and approaches evolved over time? In what ways do they continue to shape our understanding and assessment of prevention efforts today?

From the socio-historical overview provided in this article, several findings are noted:

- From its very beginnings back in the 1880s, both the purpose and practice of school-based drug education has been largely determined by the dominant "no-substance-use" injunction, which continues through the present day.
- Over this time span, opposition to the informational component of "no-use" drug education efforts has come from two markedly different directions:
 - opposition to exaggerated and/or erroneous graphic portrayals of the consequences of any substance use, often referred to as "scare tactics"; and
 - opposition by others who, viewing most informational efforts as merely advertising "forbidden fruits" to impressionable young minds, advocate Just Say Nothing instead.
- Although alcohol and tobacco were the primary targets of early Just Say No prevention campaigns, following the repeal of Prohibition in the early 1930s, they were soon replaced by many of the demonized drugs of today (i.e., heroin, cocaine, marijuana).
- Although advocacy of "informed choice"/"harm reduction" (Just Say Know) educational perspectives have always been present in some form, such efforts have only thrived for alcohol since the repeal of Prohibition, and for all drugs for a brief period during the 1970s.
- Formal evaluation of school-based drug education, begun in the late 1960s, continues to be constrained by its identification with the narrow goals of no-use programs at the expense of worthwhile alternative harm reduction approaches.

Taken together, these findings underscore two research contributions this article makes:

Insights gained from reviewing the rich history of school-based drug education in America enables us to make more informed decisions as we proceed into yet another generation of seemingly "new" prevention approaches.

This historical analysis also provides us with instructive information regarding ways of enhancing drug education evaluation.

For the purposes of this review, a comprehensive archival analysis was conducted. Historical documents ranging from the earliest temperance education textbooks and teacher's manuals to the first volumes of the *School Physiology Journal* were rigorously examined. Secondary sources (such as analyses of the temperance/prohibition movements) were also included as appropriate. Through comparative examination of this vast array of documents, similarities and differences between past and present programs and concepts emerged.

Presented findings were deemed historically relevant and meaningful only after several comparative steps were undertaken. In particular, contradictory evidence, spurious relations, and rival explanations were all taken into account throughout the analytical process (Kirk and Miller 1986; Miles and Huberman 1994). For illustrative purposes, a number of quotations from historical documents are provided that serve as exemplars for many of the principal themes of this article. Each exemplar meets the criterion of being historically representative of a particular point of view or construct through comparative analysis.

This article begins with a close look at the pivotal early years of the American drug education experience. Particular attention is devoted to the enormously influential, but largely forgotten, "scientific temperance instruction" developed and proselytized by Mary H. Hunt and the Woman's Christian Temperance Union (WCTU) in the late 1800s. For half a century, this school-based approach effectively dominated the picture and established the precedents for much of what we associate with the field of drug education today.

Of central importance to this article is the respective role and importance of information provision among the various educational perspectives that have been influential at different points along the historical timeline. In light of the growing rejection of scare tactics efforts, a more realistic informed choice educational approach is described. In marked contrast to Just Say No directives that currently dominate drug prevention strategies in the United States, this Just Say Know approach strives to minimize the potential harm resulting from the misuse of any drug, regardless of its legal status. The article concludes with an in-depth look at what an appropriate and defensible evaluative strategy might look like, one that assesses the effectiveness of such a harm-reduction approach in achieving its desired objectives. As a result of largely unquestioned no-use objectives within the field, traditional evaluations have left unexamined certain critical educational effectiveness indexes. In general, evaluations of drug education programs provide little insight as to the quality of information conveyed by such efforts or their ability to reduce potential misuse of various substances.

WCTU "SCIENTIFIC TEMPERANCE INSTRUCTION": THE ORIGINAL JUST SAY NO DRUG EDUCATION

The advent of formal school-based drug education dates back to the early 1880s. It was at this time that members of the recently formed WCTU sought to take preventive action against alcohol, tobacco, and other intoxicating "narcotics" by reaching youth before they had begun to use them. The WCTU perceived alcohol and, to a lesser extent, tobacco, opium, and other narcotics to be the primary cause for most of the social ills besetting a rapidly growing and changing America (Austin 1895; Cook 1895).

In 1880, the WCTU established a Department of Scientific Instruction to oversee their temperance education efforts, with Mary H. Hunt at the helm (Hunt 1891; Leiter 1890). Francis Willard, the charismatic leader of the WCTU during this time, described the passion that drove her fellow crusader, citing Mrs. Hunt's fervent belief that the success of the temperance reform depends on the universal education of the successive generations of the people as to the real nature and physiological effects of alcoholic beverages. To accomplish this in the United States, she now devoted her life (Willard 1886, 252-53).

For the next quarter of a century up until her death in 1906, Mrs. Hunt would devote extensive time and energy to ensuring the implementation of carefully planned "Scientific Temperance Instruction" in schools across the nation. Properly taught, she believed that students not only would be abstainers for life but also would bring a school-bred bias in favor of prohibition to the ballot box when they were old enough to vote (Erickson 1988; Lender and Martin 1987; McCarthy 1964).

Mrs. Hunt and her colleagues were enormously successful in promoting their agenda on a number of legislative fronts. Bolstered by continuous support from the National Education Association among other groups, the WCTU's campaign to enact scientific temperance education laws met with rapid acceptance (Mezvinsky 1961). Beginning with Vermont in 1882, within a decade, all but 10 states had compulsory temperance instruction laws. In 1886, the federal government mandated such education in all nationally owned schools, including the military academies (Hunt 1897).

Reflective of similar statutes enacted elsewhere, the California Education Code of 1887 provided the following directions in regard to curriculum:

Instruction must be given in all grades of school and in all classes during the entire school course, in manners and morals, and upon the nature of alcohol and narcotics and their effects upon the human system. (Hunt 1891, 72)

By 1901, every state and territory had passed temperance education laws, resulting in more than 22 million students receiving compulsory instruction on the evils of alcohol, tobacco, and other narcotics during the following school year (Mezvinsky 1961).¹

Mrs. Hunt and her colleagues also devoted considerable time and effort to conducting teacher training as well as assisting authors of popular school physiology and hygiene texts in making the necessary revisions required to earn coveted WCTU approval (cf. prefaces to Blaisdell 1893; Brown 1887). To be deemed "acceptable," textbooks were required, among other things, to teach that "alcohol was a dangerous and seductive poison," advocate total abstinence, and avoid all references regarding alcohol as a medicine (Leiter 1890; Hunt 1897). If the texts did not receive the approval, certain states and schools were reluctant or even forbidden from using them (Hunt 1891).

Within four years after the textbook revision program had begun in 1888, the number of endorsed textbooks had risen from 0 to 23. Still more were to follow in subsequent years as authors and publishers sought to get approval to join in the lucrative market created by the rapid spread of mandatory temperance instruction laws across the nation (Mezvinsky 1961).

Driven by their fervent commitment to stamping out "Demon Alcohol," tobacco, and other narcotics, the WCTU resorted to what is often derisively referred to as scare tactics drug education. These tactics took various forms: graphic portrayals of the physical harms and moral degradation associated with substance use, the message that all use—no matter the amount or form—was abuse, and the now familiar refrain that the latest scientific findings backed their claims.

The bulk of narcotics education was devoted to graphic depictions of their inherent dangers (Blaisdell 1893; Brown 1887; Nattress 1893; Overton 1897). Even a cursory review of such popular school texts as Dr. Blaisdell's *Our Bodies and How We Live* (1893), Dr. Brown's *The House I Live In* (1887) or Dr. Overton's *Applied Physiology* (1897) reveals the remarkable attention devoted to depicting the dreadful consequences awaiting youthful experimentation. Reflective of the comprehensive approach in which most of these temperance texts discussed intoxicant dangers, *Public School Physiology and Temperance* (authorized by the Education Department of Ontario, Canada in 1893) portrays the dangers posed by "alcoholic stimulants and narcotics" toward every bodily organ or function explored in the manual. As the physician/author explains in his preface:

The pupil is, in this way, at every turn confronted with the evil effects of alcohol and tobacco, the dangers accompanying their use, and the tremendous risk of tampering with such powerful agents of destruction. (Nattress 1893, iii)

The WCTU emphasized that students must be informed of the "appalling effects of drinking habits upon the citizenship of the nation, the degradation and crime resulting" (Hunt 1897, 47). Not surprisingly, each of the approved temperance texts devotes considerable detail to describing the alarming effects of alcohol and other intoxicants on the moral centers of the brain. In the Blaisdell text, it is stated that "Beer is responsible for many crimes. It seems to have a benumbing effect upon the moral nature that prepares the drinker for wicked and cruel deeds and for deliberate crime" (Blaisdell 1893, 89).

The WCTU-approved hygiene texts took great pains in pointing out the inherent risk accompanying the use of alcohol in any form. As one of the "fundamental facts" underlying the "Science of Total Abstinence," Mary Hunt (1902) maintained that

alcohol, like all other narcotics, has the power when taken frequently, even in small quantities, to create an abnormal desire for more, which may become uncontrollable and destructive. (P. 2)

Because as this group believed, "alcohol, like all other narcotics has the power . . . to create an abnormal desire for more," each of the WCTU-approved texts took great pains in repeatedly pointing out the inherent risk accompanying the use of alcohol in any form. Exemplifying this absolute no-use injunction, Blaisdell (1893) rejected the popular notion that consumption of low alcohol products such as hard cider, beer, and light wine is somehow safer than that of hard liquor, for the simple reason that

no one is safe who begins to take any liquor containing alcohol. Entire abstinence is the only safeguard against forming the alcoholic appetite . . . no liquor containing alcohol should ever be used as a flavoring for pies, puddings-sauces, jellies, or any other article of food. (P. 87)

From this standpoint, it is easy to understand the WCTU's position that "The children of this country must not be sacrificed to false teachings in favor of moderate drinking" (Hunt 1904, 17). In a related fashion, the approved textbooks describe how the use of any narcotic fosters not only an "uncontrollable appetite" for more of the same substance, but often induces one to experiment with other intoxicants as well. These approaches linking any substance use with abuse provided one of the earliest definitions of a "no-substance-use" message.

In this early version of the "stepping stone" theory, cigarettes were often accorded a prominent role (Sullum 1996). After alcohol, which commanded the lion's share of concern in the various texts, tobacco, particularly in the form of cigarettes, typically occupied second place in terms of the amount of space devoted to its discussion. Cigarettes were also popular subjects in the plethora of songs and poems written for children of all ages to sing or recite. A particularly intriguing verse appearing in a turn of the century manual called *Temperance Helps for Primary Teachers*, anticipated a phrase that would become popular eight decades into the future:

Say No! to tobacco, that poisonous weed.
 Say No! to all evils, they only can lead
 To shame and to sorrow; Oh, shun them, my boy,
 For wisdom's fair pathway of peace and of joy.
 (Freese 1901, 55)

Underscoring his assertion that "the use of cigarettes by young people cannot be too severely condemned," Blaisdell quotes an editorial that appeared in the *New York Medical Record*:

The evils of tobacco are intensified a hundred fold upon the young. Here it is unqualifiably and uniformly injurious. It stunts the growth, poisons the heart, impairs the mental powers, and cripples the individual in every way. . . . Sewer gas is bad enough, but a boy had better learn his Latin over a man-trap than get the habit of smoking cigarettes. (In Blaisdell 1893, 203)

In reading Mrs. Hunt's monthly *School Physiology Journal* as well as temperance texts targeted for the advanced grades, one finds copious quotations from prominent sources such as the above to provided credence and authority to sundry claims. Countering the popular public sentiment of the time that beer was considerably less dangerous than the hard liquors as a result of its much lower alcohol content, Blaisdell (1893) provides a detailed refutation for why this simply was not the case. Relying on the best of authorities for support, he explains how

a copious beer-drinker often looks the very picture of health, and boasts of the healthfulness of his favorite beverage; but the testimony of physicians, surgeons, and life-insurance companies is that the beer-drinker, of all others, is most liable to swift and sudden death from some slight causes. The surgeon

dreads him for a subject, for his blood is often in such a state that a slight cut may develop into a gangrenous wound that ends quickly in death. A slight cold brings on a fatal pneumonia in spite of the best physician's efforts. (Pp. 89-90)

From early on, scientific temperance instruction and the information conveyed in the WCTU-approved textbooks were the subject of severe criticism. In fact, opposition came from many fronts, ranging from the presidents of several prominent universities to the well-respected "Committee of Fifty to Investigate the Liquor Problem," a group of scholars organized to study the alcohol issue in the 1890s (Bowditch and Hodge 1903). Articles and editorials also appeared in the popular press, taking both the instruction and the endorsed textbooks repeatedly to task for exaggeration and attempting to bring about moral reform under the guise of science (Mezvinsky 1961; "Editorial" 1887). Although most of these critics acknowledged the considerable costs caused by drunkenness and alcoholism to both individuals and society, they generally pointed to the responsible moderation practiced by the majority of American drinkers as posing little or no harm.

It was the Physiological Sub-Committee of the Committee of Fifty that waged the most concerted attack on the activities of Mary Hunt and her associates at the WCTU. Based on their extensive review of "scientific temperance instruction" as it existed at the turn of the century, the subcommittee concluded that it was "neither scientific, nor temperate, nor instructive" (Bowditch and Hodge 1903, 44).

In particular, the subcommittee argued that the absolutist nature of such instruction had failed

to observe the distinction between the diametrically opposite conceptions of use and abuse. . . . It should not be taught that the drinking of one or two glasses of beer or wine by a grown-up person is very dangerous, for it is not true. (Bowditch and Hodge 1903, 44, xxii)

Outraged by these allegations, Mary Hunt organized a full-scale counter-attack, even going so far as obtaining a resolution passed by the U.S. Senate in support of her approach. More than 100,000 copies of her *Reply to the Physiological Sub-Committee of the Committee of Fifty* were published, accusing them of bias and misrepresentation of facts (Hunt 1904).

Both sides repeatedly invoked the latest findings of "science" to justify their respective positions on school-based drug education. This never-ending battle for ownership of the "truth" is equally commonplace in contemporary

debates between roughly similar oppositional camps today. In regard to information provision about illicit drugs, one finds accusations of scare tactics coming from one direction and "advocating use" issuing from the other.

After excoriating scientific temperance instruction on many levels, the Physiological Sub-Committee lamented how entrenched in the educational system and the minds of legislators this total abstinence approach had become. As a consequence, they conceded that

the removal of this educational excrescence will be no easy task . . . a prolonged struggle will be necessary to free our public school system from the incubus which rests upon it. (Bowditch and Hodge 1903, 45)

History indeed proved their assessment to be correct, as scientific temperance instruction would continue to thrive for another three decades. Always alert to signs of its diffusion overseas, Mary Hunt enthusiastically noted the recent enactment of laws in France that mandated anti-alcohol instruction for students at all grade levels (Hunt 1905).

In actuality, compulsory education on alcohol in France, which had begun in 1895 in response to the rapid rise in distilled spirits consumption, corresponded much more with the sentiments of the Committee of Fifty and other groups opposed to the WCTU's scientific temperance instruction. Despite similarities in the methods used to convey their intended messages, French students were being provided with a very different injunction than their American counterparts. They were only encouraged to abstain from the use of distilled liquors. When it came to fermented beverages such as wine, cider, and beer, students were instructed to "use but don't abuse" and strive to drink in moderation (Gershman 1987). As such, the French system of temperance instruction appears to represent the first attempt by a government to enact compulsory "responsible use" or harm-reduction drug education.

Interestingly enough, the American temperance movement had originally organized around a similar injunction against distilled spirits rather than all alcoholic beverages in its formative years earlier in the 19th Century. However, sentiment had overwhelmingly moved against the use of all alcoholic beverages by the time the WCTU and scientific temperance instruction came on the scene. (Lender and Martin 1987; Levine 1978; Rorabaugh 1979)

In the eyes of many, the enactment of national Prohibition signaled that the WCTU's efforts had passed the ultimate evaluative test it had set for itself. As Mary Hunt had confidently predicted back in 1906:

The child is born who will see the last legalized saloon, brewery and distillery of alcoholic drinks go from the United States, if the people now enforce the temperance education laws they have enacted. (Hunt 1906, 67)

In the WCTU's eyes, more than 35 years of educating the "abstainers of tomorrow" had led to this reward. When national Prohibition went into effect in the 1920s, however, the WCTU responded by pursuing their temperance education efforts with even greater fervor. Aided by the National Education Association, they sought to guard against any complacency that might threaten their hard-fought success (McCarthy 1964; Milgram 1976; "Teaching Topics" 1931; Ormond 1929).

POST-PROHIBITION "NARCOTIC EDUCATION": THE QUIET YEARS OF JUST SAY NOTHING

The repeal of the "Noble Experiment" in the early 1930s signaled the beginning of the end for scientific temperance instruction. Despite the best efforts of its advocates, it soon began to recede from the public consciousness and school curriculum. States and local school districts where "dry" opinion remained strong continued to instruct students about the evils of alcohol and the benefits of abstinence. The WCTU, NEA, American Medical Association, and other groups helped to ensure that acceptable textbooks were readily available to assist them (cf. Palmer 1937).

Amid the societal ambivalence toward alcohol and tobacco observed during the middle years of this century, it was not unusual to find many schools providing little or no instruction at all about these substances (Lender and Martin 1987; Milgram 1976).

Across the nation, however, the growing acceptance of the WCTU's public enemies #1 and #2 among the adult population led to the increasing abandonment of total abstinence instruction regarding alcohol and tobacco in schools. Increasingly, some schools began providing a form of "responsible decision-making" education about alcohol, described by not solely focusing on no-use and scare-tactic oriented programs, while at the same time providing more objective information than in the past. This was typically coupled with an injunction for students to await legal drinking age before putting their newly gained wisdom to the test (Milgram 1976). As the opening paragraphs of a popular 1952 pamphlet admonishes the young reader, "Wait until you are grown up!" is the emphatic advice that science and society offer to all young

people when they consider beginning to smoke or to drink (Rathbone 1952, 1).

Interestingly enough, the same author, a professor of health and physical education at Columbia University, concludes her remarkably tempered discussion of both the pros and cons of alcohol use by asking:

If you are a boy or girl of school age, should you begin drinking? If you have already begun, should you continue? This is a choice you must make for yourself. It is hoped that the information and ideas presented above have helped you to arrive at an intelligent decision. (Rathbone 1952, 39)

During the next several decades up through the present day, school-based alcohol education has undergone a number of shifts in dominance between these three competing perspectives of Just Say No, Just Say Know or Just Say Nothing (Langton 1991; Milgram 1976).

In regard to the WCTU's other targeted "narcotics" (i.e., opiates and cocaine), a Just Say Nothing perspective eventually supplanted Just Say No efforts for a time in the decades following their prohibition. Unlike alcohol, an informed choice approach was rarely, if ever, broached for these substances until the 1970s. This divergence in accepted approaches for different intoxicants is explained by Rathbone, who observes that

with tobacco, the main problem is not to go beyond moderation. . . . With regard to alcohol, a person may also indulge moderately with no very serious effect. . . . With narcotics, the situation is different. The cumulative effects are extremely rapid, and, within a matter of days after taking the first marihuana smoke or "shot" of pain-killing drug, one may be an addict. (Rathbone 1952, 3-4)

At least initially, the increasing demonization of heroin and cocaine in the 1920s and marijuana in the 1930s was accompanied by efforts to alert the populace to these menaces threatening the nation. Particularly prominent among these were the founding of the International Narcotic Education Association by war hero, former Alabama congressman, and ardent prohibitionist, Richard P. Hobson. Among its other activities, Hobson's organization staged a number of international Narcotic Education Weeks for both school children and the general public from the mid-1920s through the early 1930s (Morgan 1981).

Although initially popular, enthusiasm and monetary support for such efforts rapidly declined, largely as a consequence of active opposition of

prominent government officials and other influential parties. In general, these individuals viewed control of the "narcotics problem" as best left up to the legal profession, who appeared to have the problem well in hand. Many critics of Hobson and others who sought to alert children to the dangers posed by narcotics viewed such efforts as merely advertising the hitherto unknown attractions of these "forbidden fruits" to impressionable young minds (Anslinger and Tompkins 1953; Morgan 1981; National Commission on Marijuana and Drug Abuse 1973). In one prominent book written just before the repeal of Prohibition, the author summed up the dilemma facing those who would wish to provide school-based "narcotic education":

The proposed introduction of narcotic education into the public schools, like sex education, raises some questions which it is not easy to answer. One of the most serious of these is the reputed danger of stimulating the curiosity and adventure interest of the child through emphasizing either negatively or positively the unusual effects of drugs upon both mind and body. . . . Whether or not such instruction would actually have these effects has never been ascertained by truly scientific investigations. Theoretically, however, the danger deserves serious consideration. It leads us to suggest that the more indirect methods of education may serve the purposes intended. (Payne 1931, 219-20)

As evidenced in the above quotation, the 50-year legacy of compulsory school-based scientific temperance instruction was already becoming forgotten history among the new drug warriors emerging in the post-prohibition era. Despite the WCTU's best efforts, alcohol and tobacco were no longer considered part of the "narcotic problem," let alone the primary representatives of it.

Foremost among those arguing against school-based drug education was Harry Anslinger, who was to play a dominant role in all matters of drug policy and control during his long-standing reign as commissioner of the Federal Bureau of Narcotics (now Drug Enforcement Administration) from 1930 through 1962. Through his efforts, education of the populace remained largely confined to well-orchestrated media blitzes for a number of decades (Beck 1988; Zinberg and Robertson 1972).

A particularly notorious example of this was the time Anslinger began to issue frequent press releases in 1935 that documented the horrible crimes committed by marijuana-intoxicated youth and/or addicts. With headlines announcing "The New Narcotic Menace" and the "Crusade Against Marijuana," articles that contained remarkably similar accounts appeared in major newspapers and national magazines. This was hardly surprising, in that the

mainstream media relied almost solely on the Federal Bureau of Narcotics (FBN) for their facts, figures, and requisite horror stories (Becker 1963). The success of Anslinger's efforts in this particular instance was found in the smooth passage of the Marijuana Tax Act of 1937 by Congress (Dickson 1968; Morgan 1981).

Anslinger's media campaigns were primarily focused on influencing public opinion to garner necessary support for his own personal bureaucratic objectives. Throughout his lengthy tenure, he actively discouraged most efforts at educating youth about illicit drugs. Regardless of the "fear quotient" conveyed by any particular approach, Anslinger contended that intensive or informational forms of school-based drug education merely served to arouse unnecessary curiosity among impressionable youth (Anslinger and Tompkins 1953; Finlator 1973). Nevertheless, as Wallack (1980) observes:

The efforts of the FBN in the 1930's, in what could probably roughly be characterized as the first federally sponsored drug education campaign, established a trend that was to be followed through the 1960's—the use of sensationalism and scare tactics and the avoidance, repression, or minimization of scientific information. (P. 57)

RESPONDING TO THE 1960s DRUG CRISIS: THE RETURN OF SCARE TACTICS EDUCATION AND ITS DISCONTENTS

In 1963, an "Advisory Commission on Narcotic and Drug Abuse" appointed by President Kennedy recommended a fundamental shift in direction regarding drug abuse prevention. In advocating much more intensive school-based drug education, they challenged the long-dominant view held by many, including Anslinger, who had only recently stepped down and passed away the previous year from his position as head of the Federal Bureau of Narcotics:

There is a vigorous school of thought which opposes educating teenagers on the dangers of drug abuse. The argument runs that education on the dangers of drug abuse will only lead teenagers to experimentation and ultimately to addiction. The Commission rejects this view. . . . The Commission feels that the real question is not whether the teenager should be educated, but who should educate him? Should it be the street corner addict, or should it be the

schools, churches, and the community organizations? (The President's Advisory Commission on Narcotic and Drug Abuse 1963, 18)

Perhaps somehow anticipating the unprecedented explosion in youthful drug use that would soon sweep the nation, the committee recommended more diligent provision of information-based "fear-arousal" drug education programs for student populations:

The teenager should be made conscious of the full range of harmful effects, physical and psychological, that narcotics and dangerous drugs can produce. He should be made aware that, although the use of a drug may be a temporary means of escape from the world about him, in the long run these drugs will destroy him and all he aspires to. (The President's Advisory Commission on Narcotic and Drug Abuse 1963, 17-18)

In response to the mounting youthful drug crisis confronting the country, President Nixon followed up his declaration of a "War on Drugs" with a mandate to help ensure that all students from kindergarten through high school would receive the education considered vital to stemming the threatening epidemic.

Combined with already existing state and local funding, this massive influx of federal support only added to the confusing array of haphazardly conceived prevention efforts seeking to convince youths to stay away from drugs such as LSD and marijuana embraced by the counterculture. The vast majority of drug education programs were not well coordinated or conceived (National Commission on Marihuana and Drug Abuse 1973).

Added to this muddle was the excessive reliance on fear arousal or scare tactics approaches characteristic of the majority of school-based prevention efforts. In one of the first comprehensive explorations of student drug use, Blum and Associates (1969) lamented the "increasing polarity" and distrust experienced by students in response to the punitive and moralistic paternalism shown them. As a consequence,

Dishonesty in this area weakens credibility in all areas; hypocrisy generates wide distrust; reliance on external control and authoritative pronouncement weakens the development of internal controls and learning to make informed decisions. . . . Educators are in the uncomfortable position of knowing that most prevalent methods of drug education are ineffective and in many cases contribute to the very problem they seek to control. (Blum et al. 1969, 356-57)

Surveying the chaotic prevention field in existence at the time of their review in the early 1970s, the congressionally appointed National Commission on Marihuana and Drug Abuse arrived at essentially similar conclusions, observing that "no drug education programs in this country, or elsewhere, has proven sufficiently successful to warrant our recommending it" (National Commission on Marihuana and Drug Abuse 1973, 357).

Concerned about the profound disorganization of the field and the troubling implications accompanying overreliance on scare tactics approaches, the commission further concluded that "the avalanche of drug education in recent years has been counterproductive" and called for a national moratorium on drug education until better strategies for implementation and evaluation could take place (National Commission on Marihuana and Drug Abuse 1973).

HARM REDUCTION DRUG EDUCATION IN THE 1970s: TEMPORARY LEGITIMATION OF INFORMED CHOICE/RESPONSIBLE DECISION-MAKING APPROACHES

Buoyed by mounting frustration with traditional drug education efforts, a number of promising developments began to emerge in the 1970s. A number of these were noted in a series of progress reviews in the prevention field put out by the federal government (Boldt, Reilly, and Haberman 1973; DuPont, Goldstein, and O'Donnell 1979; National Institute on Drug Abuse 1975; Nellis 1972). These reviews, written and published by the federal government's National Institute on Drug Abuse (NIDA) and its predecessor the National Institute of Mental Health (NIMH) indicate a shift toward a pragmatic stance befitting the realities of the burgeoning drug scene. The goal of this new stance—misuse or abuse prevention—was fundamentally different from the goal of the past stance—total abstinence (Edwards 1973; Feinglass 1972; McCune 1973; Richards 1972; Swisher 1979). This goal of minimizing or preventing problematic consequences associated with substance use assumes a harm-reduction perspective rather than an abstinence-based perspective. At the time, these programs were not labeled harm-reduction programs; that term was almost unknown in the drug abuse field until the advent of the AIDS epidemic in the 1980s.

In rejecting abstinence-only approaches, Australian researcher Marion Watson describes the pragmatism underlying such efforts:

Harm reduction in relation to drug use is the philosophical and practical development of strategies so that the outcomes of drug use are as safe as is situationally possible. It involves the provision of actual information, resources, education, skills and the development of attitude change, in order that the consequences of drug use for the users, the community and the culture have minimal negative impact. (Watson 1991, 14)

Exemplifying the profound differences in program philosophy and objectives between past and emerging models, Swisher (1979) listed some of the key assumptions that were increasingly guiding drug education during the latter half of the 1970s. Among these guiding principles, two of particular significance were the following:

- A reasonable goal for drug-abuse prevention should be to educate for responsible decision making regarding the use of all drugs (licit and illicit) for all ages.
- Responsible decisions regarding personal use of drugs should result in fewer negative consequences for the individual (Swisher 1979, 427).

In commentary accompanying a comprehensive overview of the prevention field in the 1970s, the editors of the federal government-sponsored *Handbook on Drug Abuse* observed how Swisher's recommended abandonment of abstinence-only approaches in favor of harm-reduction approaches was in agreement with specific recommendations made by two federal government panels convened on prevention issues in the late 1970s. They also noted how Swisher's views concurred with a "wide, highly competent, and influential segment of the drug community and of the wider professional community" (DuPont, Goldstein, and O'Donnell 1979, 406).

The closing chapter of the *Handbook on Drug Abuse* reflects the significant shift in governmental policy that was taking place toward the end of the 1970s. It is here that lead editor and former NIDA director, Robert DuPont (1979), provides his vision of a more tolerant, humane "Future of Drug Abuse Prevention." In this essay, he declares:

We as a Nation have come out of a period which extended from the 1920s into the mid-1960s. . . . We have gone through a decade of profound reaction to this ill-informed early period. Many Americans, especially the most sophisticated, have concluded that we must turn drug abuse policy issues over to the scientist to avoid ever again repeating the errors of earlier decades. (P. 451)

THE RETURN OF JUST SAY NO: FROM ILLUSORY SUCCESS IN THE 1980s TO LESSONS (RE)LEARNED IN THE 1990s

After making the bold pronouncement quoted above, DuPont enthusiastically hopped on board the "Parent Power" bandwagon the following year. The parent power movement, which rapidly gained strength and influenced government policy during the Reagan/Bush era, was a group of activist parents who began organizing in Georgia in the late 1970s over concern about their children's use of drugs. Writing in early 1980, DuPont explained the unexpectedly rapid reascendancy of a form of drug education that he had previously dismissed as the "errors of earlier decades":

A few years ago parent power's simple message—"no"—on drug use by youths contrasted with the prevailing message of the experts—"Your body is your own; it is your decision to use or not to use any drug." However, the experts have seemed relieved to have parents tell them the bottom line on drug abuse prevention and have generally set about finding, within their own areas of expertise, constructive ways to join the parent power movement. (P. 2)

The parents' power movement rapidly transformed prevention policy and practice across the country; they effectively collaborated with government and professionals in the substance abuse field in carrying out the necessary revisions of curriculums and materials to reflect the strict adherence to no-use messages and "zero tolerance" in our latest "War on Drugs" (Baum 1996; Manatt 1979; U.S. Department of Education 1988; White House Conference for a Drug Free America 1988).

Since the beginning of the parent power movement, the steady growth of prevention funding have helped to ensure increasing sophistication and wider implementation of prevention efforts. In popular school-based drug education programs, students are often given information by police officers and former drug addicts, told of the dangers of using drugs through techniques often intended to arouse their fears, given rewards such as T-shirts and bumper stickers for complying with the no-use message, and taught how to refuse the use of substances through decision-making programs (Brown, D'Emidio-Caston, and Pollard 1997).

However, the rigidity of no-use dictates of programs in the 1980s and 1990s practically ensures an incomplete and biased presentation of the current knowledge regarding both legal and illegal drugs. In the process, the

"decision-making skills" approach that came into fruition back in the 1970s has been essentially co-opted by contemporary efforts that no longer permit students to make responsible, informed choices regarding their drug use. Instead, the modern version consists of refusal skills training in which any "decisions" to be made have been firmly predetermined in advance for the target population (Baum 1996; Brown, D'Emidio-Caston, and Pollard 1997; Duncan 1992; Rosenbaum 1996).

In addition to the school-based programs, a massive media campaign led by the Partnership for a Drug-Free America is attempted to prevent kids from using drugs through television, magazine, and newspaper ads. These ads, which often feature scare tactics, do not tackle the issues of alcohol, tobacco, or pharmaceutical. This notable absence may be partially explained by the fact that Anheuser-Busch, Philip Morris, RJR Reynolds, Hoffman-LaRoche, and SmithKline Beecham are some of the companies who have generously bankrolled the slick and often controversial Partnership ads attacking non-sanctioned drugs not manufactured by them (Baum 1996; Blow 1991; Cotts 1992; McShane 1992; Miller 1996; Reeves and Cambell 1994).

The continued rise in youthful drug use during the 1990s has done much to temper the considerable optimism among Just Say No approaches, which had accumulated as a result of steady declines in use observed during the previous decade. Johnston and colleagues, who conduct the annual federally funded Monitoring the Future survey of students across the country, account for the contemporary resurgence by noting that

we may be seeing the beginning of a turnaround in the drug abuse situation more generally among our youngest cohorts—perhaps because they have not had the same opportunities for vicarious learning from the adverse drug experiences of people around them and people they learn about through the media. Clearly, there is a danger that, as the drug epidemic has subsided considerably, newer cohorts have far less opportunity to learn through informal means about the dangers of drugs. This may mean that the nation must redouble its efforts to be sure that they learn these lessons through more formal means—from schools, parents, and focused messages in the media, for example—and that this more formalized prevention effort become institutionalized so that it will endure for the long term. (Johnston, O'Malley, and Bachman 1994, 24)

Although the call for more comprehensive and consistent drug education in American schools and society is appealing, what is needed is something qualitatively different from the status quo. With few exceptions, comprehensive evaluations and metaevaluations of current prevention efforts have

generally revealed nonexistent or negligible effects in affecting illicit drug use among target populations (Bangert-Drowns 1988; U.S. General Accounting Office 1991; Gerstein and Green 1993; Gorman 1995; Hansen 1992; Klitzner 1987; Moskowitz 1989; Schaps et al. 1981; Tobler 1986).

As in the past, the moral certainties driving the current drug war continue to take precedence over objective niceties, ensuring that truth is once again a frequent casualty in drug "education" campaigns. The disconcerting observation made a quarter of a century ago is all too apt in describing our current state of affairs: "We have become so convinced of the nobility of our objectives that we easily rationalize our deceit and dishonesty" (Fulton 1972, 37).

If this is indeed the case, then simply "redoubling" what we currently call "drug education" as a means of correcting youthful ignorance about the dangers posed by nonsanctioned drugs is more likely to prove counterproductive, as suggested by the 1960s experience. We should remain deeply concerned about the quality and accuracy of information conveyed within the current Just Say No climate. Unfortunately, for far too long, drug education has been "little more than a frantic search for the best method for persuading youths to abstain" (Chng 1981, 14). So long as this state of affairs continues, one can anticipate continued overreliance on what youthful target audiences dismiss as scare tactics propaganda. Ironically, the resulting "credibility gap" aptly illustrates how such efforts can indeed "send the wrong message," ultimately fostering widespread distrust and discounting of all messages—no matter how credible.

EVALUATING INFORMATIONAL DRUG EDUCATION APPROACHES: WHAT HAVE WE LEARNED FROM HISTORY?

The need for evaluation of school-based drug education was first voiced by the Physiological Subcommittee of the Committee of Fifty back at the turn of the century (Bowditch and Hodge 1903, 43). In answer to their query as to whether the WCTU possessed "any data from any state showing decrease in consumption of alcoholic drinks since the passage of temperance laws," Mary Hunt responded by citing an increase in longevity among the American population observed during this time. She also went on to note that the gain in per capita use of alcoholic liquors throughout the country during the past 11 years was only a third as great as the previous 11-year period when scientific temperance instruction was being first introduced. Finally, she cited the widespread recognition by other countries that the increasingly abstinent

American workplace was largely responsible for the phenomenal success of the nation. In concluding her "evaluation," Hunt bluntly states "If the Subcommittee deny that this education has been a factor in securing the above results, here stand the facts" (Hunt 1904, 9).

In reality, little formal evaluation of drug education was done until the 1970s (Boldt, Reilly, and Haberman 1973; de Lone 1972; Richards 1972). Despite calls for formal evaluations voiced by government and other entities since the advent of the youthful drug explosion in the 1960s, funding and leadership in promoting such efforts was lacking. In fact, the first systematic guidelines for evaluating drug education programs did not appear until 1973 and were sponsored by the privately supported Drug Abuse Council (Abrams, Garfield, and Swisher 1973).

Evaluations of cognitive or informational drug education programs typically revealed significant gains in knowledge among program participants. However, such efforts were almost invariably found to be ineffective in either reducing illicit drug use or generating "better" attitudes toward unsanctioned substances (Goodstadt 1980; Schaps et al. 1981).

In a few notable instances, evaluative studies even revealed modest increases in experimentation following the 1970s drug education programs. In one of the first in-depth evaluations of a particular harm reduction program, Stuart unexpectedly found the only changes between the experimental and control groups to be slightly significant increases in the use of LSD and marijuana among those hearing the informational presentation. Stuart attributed these modest effects to curiosity and/or allayed fears resulting from the demythologizing of these two particularly demonized substances (Stuart 1974).

Such evaluation findings led some in the prevention field to call into question the belief that information provision is a necessary component of substance abuse prevention. Recalling similar concerns voiced by Anslinger and other authorities in the past, information-based approaches have been repeatedly taken to task and dismissed for arousing curiosity among youths and introducing them to new and stronger ways of getting high (Stickgold and Brovar 1978; Stuart 1974).

Despite the continued failure of current Just Say No programs, evaluators and other researchers in the substance abuse field remain highly susceptible to governmental pressure to avoid evaluating current policies and practices under the guise of threatening the "united front" deemed necessary to "win the war" (Brown and Horowitz 1993; Zinberg 1984).

The above concerns notwithstanding, support for the abstinence-only edict has not been without criticism from within the government itself as

evidenced by the sharp critique issued by the General Accounting Office (GAO) in its evaluation of mandated prevention objectives during the Reagan-Bush Administrations (U.S. General Accounting Office 1991). Calling into question the rigid adherence to abstinence-only approaches and other dogmatic dictates, the GAO report focused on what it believed to be the urgent need for research evaluating the efficacy of what were argued to be more realistic "Responsible use approaches (which) . . . may try to reduce the riskiest forms of use (such as drinking and driving) or encourage current substance users to cut down." (U.S. General Accounting Office 1991).

The analysis of the GAO, and America's history of drug prevention education, call for the formulation and evaluation of new harm-reduction strategies. In this regard, much can be learned from careful study of harm-reduction efforts carried out in the United States during the 1970s as well as similar approaches currently gaining in popularity overseas (Advisory Council on the Misuse of Drugs 1993; Bagnall and Plant 1987; Beck 1980; Clements, Cohen, and Kay 1990; Cohen 1989, 1993; DeHaes 1987; Duncan et al. 1994; Miller 1975).

HARM REDUCTION OR INFORMED CHOICE DRUG EDUCATION: DESCRIPTION OF AN APPROACH AND UNIQUE EVALUATIVE STRATEGIES

A number of harm-reduction drug education approaches have emerged as worthy candidates to choose from during the past three decades. Because drug issues often involve highly emotional individual feelings, group interactions, and charged social contexts, researchers have argued that prevention efforts must take effective domains and environmental conditions into account (Dembo 1979). Although the focus of this article is on cognitive drug education, it is important to note that attempts are underway to create multidimensional programs and accompanying evaluation models which use what we know about the inextricable links between feeling, thinking, and learning (Horowitz and Brown 1996; Skager and Brown forthcoming).

One cognitive harm-reduction program that will be the focus for the remainder of this analysis is the Drug Consumer Safety Education program developed by Mark Miller and staff at the University of Oregon Drug Information Center in the 1970s (Miller 1975). This approach provides age-appropriate instruction in conveying thorough drug knowledge and discrimination skills. In so doing, this program emphasizes the importance of each individual becoming an informed, analytic consumer of all substances.

ranging from prescription, over-the-counter, industrial chemicals and herbals, as well as the usual legal and illegal drugs specifically targeted by traditional drug education (Beck 1980; Burbank and Miller 1995; Miller 1975; Miller and Burbank 1997). In addition to providing information on the effects of various substances and the ways to prevent harm if experimenting with substances, the program provided information regarding referrals for treatment to those in need.

With a well-documented history revealing that drug education has "failed" if its primary or only objective is to reduce illicit drug use, it becomes important to consider whether more modest and realistic goals are attainable and desirable. In contrast to traditional abstinence-oriented approaches, the Drug Consumer Safety program strives to reduce potential harm resulting from the unintentional misuse of various substances. In this conception, the term *misuse* refers to instances in which individuals may be susceptible to potential harm as a result of inadequate knowledge about the substance(s) they are using, whether licit or illicit. From this framework, drug misuse can be seen as a major problem affecting all parts of society, not just youth. The vast majority of people go through life with little understanding of the powerful and complex substances they use for a variety of recreational and palliative purposes (Beck 1980; Miller 1975).

Evaluation of harm-reduction education approaches such as that described above calls for innovative research design and instrumentation. Accomplishing such an endeavor requires abandoning the unrealistic assumptions and rigid precepts that have effectively constrained acceptable prevention research and practice since the early 1980s (Brown and Horowitz 1993; Clements, Cohen, and Kay 1990; Dorn and Murji 1992; Duncan 1992; Goodstadt 1989; Marlatt and Tapert 1993; Moore and Saunders 1991; Rosenbaum 1996; Worden 1979). The development of a proper evaluation technique is necessary for convincingly devising a means to assess the overall efficacy of such programs in reducing potential (or real) drug misuse. This is understandably a more modest but, ultimately, more realistic goal for short-term school-based drug education than the traditional objective of reducing use (Beck 1980, 1986).

An essential first task involved in such an undertaking would be the development of an instrument that strives to assess potential drug misuse among target populations. As such, the instrument would necessarily include both a survey gauging frequency of use for various substances and an extensive drug knowledge questionnaire. I am not aware of any drug knowledge questionnaires employed for evaluative purposes that have been specifi-

cally designed to assess potential misuse. To remedy this, the knowledge questions utilized in this questionnaire (to be answered both pre- and post-intervention) would be carefully chosen for their ability to gauge potentially serious drug misuse problems. That is, if a user (or prospective user) of a particular drug does not know the correct answers and/or believes incorrect answers to questions concerning that substance, this ignorance could arguably prove to be harmful at some point in time. Special attention would be focused on heavy users of particular drugs, since faulty or incomplete knowledge among these groups would predictably result in a greater incidence of drug misuse than for less frequent users or abstainers.

Given the considerable differences of opinion that exist regarding the "known" dangers or risks associated with use of various drugs (particularly illicit ones), considerable attention must be given to how best to design a knowledge questionnaire of defensible validity. To achieve this goal, a team of carefully chosen substance abuse experts might be consulted in determining the most crucial (as well as acceptable) questions to ask about each drug.

What such a drug knowledge questionnaire would hope to gauge is illustrated by the following questions pertaining to alcohol use. For instance, do frequent users of alcohol actually know how many beers or drinks it takes within a set period of time to become legally intoxicated? Are they aware of the myths and realities concerning ways of reducing alcohol intoxication? Are users cognizant of the many potentially hazardous interactions that exist between alcohol and such common over-the-counter or prescription remedies as tranquilizers, antihistamines, aspirin, or acetaminophen? Whether out of ignorance or under the influence of prevailing misconceptions, the potential for misuse and actual harm is heightened among alcohol users (or prospective users) unable to correctly answer one or more of the above questions. Similarly, the level of risk is even further accentuated among frequent or binge drinkers in the sample (Beck 1980, 1986).

CONCLUSION

A careful review of the American drug education experience during the past century allows a better appreciation of the inherently political nature and troubling implications of the Just Say No perspective that has predominated throughout this time. Perhaps most significantly, we can see how the dangers of moralistic absolutism have all too often substituted indoctrination for real education in attempting to frighten youth away from using certain condemned

drugs. As revealed in this review, although the particular substances targeted by such efforts have changed over time, the approaches employed to dissuade their use have remained remarkably the same.

As described throughout much of this article, two fundamentally opposed perspectives toward school-based drug education have vied for legitimacy during the past century. Although each stance has been characterized by any number of names, they can essentially be juxtaposed as Just Say No versus Just Say Know. The first of these conveys a strict no-use or abstinence message regarding targeted substances or activities. In contrast, the second perspective focuses on fostering informed choices or decision making within a harm- or risk-reduction framework.

Although strict abstinence or no-use approaches have been the predominant form of school-based drug education during the past century, informed choice perspectives have also been apparent, particularly in the form of alcohol education since the repeal of prohibition, the growing acceptance of such approaches in America during a brief period in the 1970s, and past and present examples overseas.

Almost from the beginning, both perspectives have invoked science to justify their respective positions on school-based drug education. This was first shown in the justifiable attack on the WCTU's "scientific temperance instruction" by a large number of respected scientists and academics in the waning years of the 19th century.

The historical lessons revealed in the 100 years of drug education since that time further implicate the shortcomings of absolutist stances, which brought no alternative or ambiguous views, regardless of scientific merit, in attempting an unambiguous no-use message. As a consequence of knowing, *a priori*, the truth about certain "bad drugs," school-based prevention easily falls prey to reliance on select and suspect scientific "facts" for the purposes of indoctrination more than true education. In addition, whereas 100 years ago, the WCTU waged a moral war against the inherent evils posed by all intoxicants, well-entrenched interests have managed to muddy up contemporary waters with mixed messages in diligently waging a selective and confusing "war on some drugs."

Among the troubling connotations of current drug education, the strict abstinence dictates mandated by the government have resulted in prevention campaigns all too frequently relying on deservedly maligned scare tactics to convey a strong no-use message. Despite ever-increasing expenditures devoted to drug abuse prevention in America, a resurgence in youthful substance use has continued relatively unabated throughout the 1990s. Once again,

American youth appear to be serving ample notice of a growing rejection of what many dismiss as Just Say No propaganda. Perhaps the most alarming casualty of this approach has been the substantial loss in credibility inevitably fostered by such "drug education." Particularly among target populations possessing considerable drug experience, reliance on disinformation should be regarded as contraindicated.

Unfortunately, the objectives of prevention efforts are all too frequently stated in a general fashion that places too much faith or hope in the ability of drug education to reduce unsanctioned substance use. In doing so, the powerful influences of family, peers, and the mass media, among other societal pressures, are virtually ignored or disregarded. A more balanced perspective regarding the potential value of drug education would reduce expectations to better acknowledge the realities of competing forces and the importance of time and normal development. Drug education programs should focus on preventing problems associated with drug use. For some students, that may mean the ability to develop decision-making skills that lead them to avoid alcohol, tobacco, and illegal drugs completely; for other students, that might mean being more careful than they previously were when they are experimenting with substances; for others, that might mean making the decision to get help for a substance abuse problem. A harm-reduction program does not advocate substance use, but it does advocate the health of youth.

Drug education should focus on the learning of decision-making skills in hopes of generating more responsible, informed consumers whose choice to use particular substances would pose less problematic potential than if they were instructed to simply Just Say No or told nothing at all. To successfully accomplish this task, school-based programs must provide accurate, age-appropriate information concerning all drugs, not just the illicit ones. Therefore, it is essential to adopt what are arguably more *modest*, but *realistic*, goals that resonate with the objectives of harm-reduction perspectives.

This article has argued for a deeper understanding of how historical development in drug education shapes current evaluation models. An important albeit often ignored consideration in drug education research is as the U.S. General Accounting Office (1991) noted: the limited effectiveness of program evaluation itself. Due to the various social historical movements described in this article, drug education evaluation research has primarily determined program success based on limited aspects of Just Say No, rather than Just Say Know types of drug education. As a consequence, evaluation researchers may have deemed affective programs failures when they were not, while at the same time capturing incomplete or irrelevant information

from Just Say No programs and deeming these findings as indexes of success. Historical evidence suggests that it is time to develop evaluation methodologies reflecting promising alternatives such as harm reduction. Such methods would capture salient information in cognitive, affective, and contextual domains. During this critical period in drug education, historical patterns provide us with the opportunity to expand our programmatic and evaluation horizons from Just Say No to a more informed Just Say Know. It is hoped that this historical evidence will be used to create positive programs and, in the process, break the cycle of condemnation toward repetition of past mistakes.

NOTE

1. For a copy of an 1882 temperance education map of the United States and territories, contact Jerome Beck at the Center for Educational Research and Development.

REFERENCES

- Abrams, A., E. Garfield, and J. Swisher, eds. 1973. *Accountability in drug education: A model for evaluation*. Washington, DC: Drug Abuse Council.
- Advisory Council on the Misuse of Drugs. 1993. *Drug education in schools: The need for new impetus*. London: Her Majesty's Stationery Office.
- Anslinger, H. J. and W. F. Tompkins. 1953. *The traffic in narcotics*. New York: Funk and Wagnalls.
- Austin, B. F., ed. 1895. *The prohibition leaders of America*. Vol. 1. St. Thomas, Ontario, Canada: Alma College.
- Bagnall, G. and M. Plant. 1987. Education on drugs and alcohol: Past disappointments and future challenges. *Health Education Research* 2:417-22.
- Bangert-Drowns, R. L. 1988. The effects of school-based substance abuse education—A meta-analysis. *Journal of Drug Education* 18:243-65.
- Baum, D. 1996. *Smoke and mirrors*. Boston: Little, Brown.
- Beck, J. 1980. Drug use trends and knowledge among students enrolled in a required health course at the University of Oregon, Winter, 1980. Unpublished honors thesis, University of Oregon, Eugene.
- . 1986. An innovative approach to evaluating drug education. Unpublished paper, University of California, Berkeley School of Public Health.
- . 1988. The determinants and consequences of mass media drug portrayals. Unpublished qualifying exam, University of California, Berkeley School of Public Health.

- Beck, J. and P. A. Morgan. 1986. Designer drug confusion: A focus on MDMA. *Journal of Drug Education* 16 (3): 287-302.
- Beck, J. and M. Rosenbaum. 1994. *Pursuit of ecstasy: The MDMA Experience*. Albany: State University of New York Press.
- Becker, H. S. 1963. *Outsiders*. New York: Macmillan.
- Blaisdell, A. F. 1893. *Our bodies and how we live: An elementary text-book of physiology and hygiene for use in schools with special reference to the effects of alcoholic drinks, tobacco and other narcotics on the bodily life*. Rev. ed. Boston: Ginn and Company.
- Blow, R. 1991. This is the truth, this is the partnership for drug-free America, this is what the partnership for drug-free America does to the truth. *Washington City Paper*, 6-12 December, 29-37.
- Blum, R. H. and Associates. 1969. *Students and drugs*. San Francisco: Jossey-Bass.
- Boldt, R. F., R. R. Reilly, and P. W. Haberman. 1973. A survey and assessment of the current status of drug-related instructional programs in secondary and elementary educational institutions. In *Drug use in America: Problem in perspective. Appendix: Volume II: Social responses to drug use*, edited by National Commission on Marihuana and Drug Abuse, 455-547. Washington, DC: U.S. Government Printing Office.
- Bowditch, E. P. and C. F. Hodge. 1903. Report on the present instruction on the physiological action of alcohol. In *The physiological aspects of the liquor problem, I*, edited by J. S. Billings. Boston: Houghton-Mifflin.
- Brown, E. F. 1887. *The house I live in: An elementary physiology for children in the public schools*. Cincinnati, OH: Van Antwerp, Bragg.
- Brown, J. H., M. D'Emidio-Caston, and J. Pollard. 1997. Students and substances: Social power in drug education. *Educational Evaluation and Policy Analysis* 19 (1): 65-82.
- Brown, J. H. and J. E. Horowitz. 1993. Deviants and deviance: Why adolescent substance use prevention programs do not work. *Evaluation Review* 17(5): 529-55.
- Burbank, S. and M. Miller. 1995. *Teach your children well: A rational guide to family drug education*. Mosier, OR: Mothers Against Misuse and Abuse.
- Chng, C. L. 1981. The goal of abstinence: Implications for drug education. *Journal of Drug Education* 11 (1): 13-18.
- Clements, I., J. Cohen, and J. Kay. 1990. *Taking drugs seriously: A manual of harm reduction education on drugs*. Liverpool: Healthwise Helpline.
- Cohen, J. 1989. Drugs & AIDS prevention & education: Work with young people in the Netherlands. *The International Journal on Drug Policy* 1 (3): 19-23.
- . 1993. Achieving a reduction in drug-related harm through education. In *Psychoactive drugs and harm reduction: From faith to science*, edited by N. Heather, A. Wodak, E. Nadelman, and P. O'Hare. London: Whurr.
- Cook, J. 1895. Neal Dow's watchwords for the twentieth century. In *The prohibition leaders of America*, vol. 1, edited by B. F. Austin, 27-34. St. Thomas, Ontario, Canada: Alma College.
- Cotts, C. 1992. Hard sell in the drug war. *The Nation*, 9 March, 300-2.
- De Haes, W. 1987. Looking for effective drug education programmes: Fifteen years exploration of the effects of different drug education programmes. *Health Education Research* 2:433-50.
- de Lone, R. H. 1972. The ups and downs of drug-abuse education. *Saturday Review of Education*, 11 November, 27-32.
- Dembo, R. 1979. Substance abuse prevention programming and research: A partnership in need of improvement. *Journal of Drug Education* 9 (3): 189-208.

- Dickson, D. T. 1968. Bureaucracy and morality: An organizational perspective on a moral crusade. *Social Problems* 16:143-56.
- Dorn, N. and K. Murji. 1992. *Drug prevention: A review of the English language literature*. ISDD Research Monograph #5. London: Institute for the Study of Drug Dependence.
- Duncan, D., T. Nicholson, P. Clifford, W. Hawkins, and R. Petosa. 1994. Harm reduction: An emerging new paradigm for drug education. *Journal of Drug Education* 24 (4): 281-90.
- Duncan, D. F. 1992. Drug abuse prevention in post-legalization America: What could it look like? *The Journal of Primary Prevention* 12 (4): 317-22.
- DuPont, R. L. 1979. The future of drug abuse prevention. In *Handbook on drug abuse*, edited by R. L. DuPont, A. Goldstein, and J. O'Donnell, 447-52. Washington, DC: National Institute on Drug Abuse.
- . 1980. The future of primary prevention: Parent power. *Journal of Drug Education* 10 (1): 1-5.
- DuPont, R. L., A. Goldstein, and J. O'Donnell, eds. 1979. *Handbook on drug abuse*. Washington, DC: National Institute on Drug Abuse.
- Editorial. 1887. *Science* 13 (July): 1-2.
- Edwards, G. 1973. Perspectives on drug education. In *Drug use in America: Problem in perspective. Appendix: Volume II: Social responses to drug use*, edited by National Commission on Marihuana and Drug Abuse, 365-91. Washington, DC: U.S. Government Printing Office.
- Engs, R. C. 1979. *Responsible drug and alcohol use*. New York: Macmillan.
- Erickson, J. B. 1988. Making king alcohol tremble: The juvenile work of the Woman's Christian Temperance Union, 1874-1900. *Journal of Drug Education* 18 (4): 333-52.
- Feinglass, S. J. 1972. How to plan a drug abuse education workshop. In *Resource book for drug abuse education* 2d ed., edited by M. Nellis, 91-104. Rockville, MD: National Clearinghouse for Drug Abuse Information.
- Finlator, J. 1973. *The drugged nation: A narc's story*. New York: Simon & Schuster.
- Freese, L. M. 1901. *Temperance helps for primary teachers*. Chicago: The Woman's Christian Temperance Publishing Association.
- Fulton, G. 1972. Drug abuse education—Tell it like it is. *Journal of School Health* 72:33-7.
- Gershman, S. T. 1987. The temperance movement in the public primary schools of France. Abstract and discussion in *The social history of alcohol: Drinking and culture in modern society*, edited by S. Barrows, R. Room, and J. Verhey. Berkeley, CA: Alcohol Research Group.
- Gerstein, D. R. and L. W. Green, eds. 1993. *Preventing drug abuse: What do we know?* Washington, DC: National Academy Press.
- Goodstadt, M. 1980. School-based drug education in North America: What is wrong? What can be done? *Journal of School Health* 56:278-81.
- . 1989. Drug education: The prevention issues. *Journal of Drug Education* 19 (3): 197-208.
- Gorman, W. B. 1995. Are school-based resistance skills training programs effective in preventing alcohol misuse? *Journal of Alcohol and Drug Education* 41:74-98.
- Hansen, W. B. 1992. School-based substance abuse prevention: A review of the state of the art in curriculum, 1980-1990. *Health Education Research: Theory and Practice* 7:403-30.
- Horowitz, J. E. and J. H. Brown. 1996. Confluent education and evaluation research. In *Advances in confluent education: Integrating consciousness for human change*. Vol. 1, edited by J. H. Brown. Greenwich, CT: JAI.

- Hunt, M. E. 1891. *A history of the first decade of the Department of Scientific Temperance Instruction in Schools and Colleges, of the Woman's Christian Temperance Union*. Boston: George E. Crosby and Company.
- . 1897. *An epoch of the nineteenth century: An outline of the work for scientific temperance education in the public schools of the United States*. Boston: R. H. Foster and Company.
- . 1902. Five foundation facts underlying the science of total abstinence. *School Physiology Journal* 11(5): Extra Edition.
- . 1904. *Reply to the Physiological Sub-Committee of the Committee of Fifty*. Washington, DC: U.S. Government Printing Office.
- . 1905. Temperance instruction in schools abroad. *School Physiology Journal* 14 (5): 76.
- . 1906. Results to which this instruction has contributed. *School Physiology Journal* Extra Edition, 1 January.
- Johnston, L. D., P. M. O'Malley, and J. G. Bachman. 1994. *National survey results on drug use from The Monitoring the Future Study, 1975-1993. Volume 1: Secondary school students*. Rockville, MD: National Institute on Drug Abuse.
- Kirk, J., and M. Miller. 1986. *Reliability and validity in qualitative research*. Newbury Park, CA: Sage.
- Klitzner, M. D. 1987. An assessment of the research on school-based prevention programs. In *Report to Congress and the White House on the nature and effectiveness of federal, state, and local drug prevention/education programs*. Washington, DC: U.S. Department of Education.
- Langton, P. A. 1991. *Drug use and the alcohol dilemma*. Boston: Allyn & Bacon.
- Leiter, F. W. 1890. *Scientific temperance bulletin*. Rev. ed. Ohio Woman's Christian Temperance Union.
- Lender, M. E. and J. K. Martin. 1987. *Drinking in America: A history*. Revised and expanded edition. New York: Free Press.
- Levine, H. G. 1978. The discovery of addiction. *Journal of Studies on Alcohol* 39:143-74.
- Manatt, M. 1979. *Parents, peers and pot*. Rockville, MD: National Institute on Drug Abuse.
- Marlatt, G. A. and S. F. Tapert. 1993. Harm reduction: Reducing the risks of addictive behaviors. In *Addictive behaviors across the life span*, edited by J. S. Baer, G. A. Marlatt, and R. J. McMahon, chap. 11, 243-73. Newbury Park, CA: Sage.
- McCarthy, R. G., ed. 1964. *Alcohol education for classroom and community*. New York: McGraw-Hill.
- McCune, D. A. 1973. Drug education in California: A case study. In *Drug use in America: Problem in perspective. Appendix: Volume II: Social responses to drug use*, edited by National Commission on Marihuana and Drug Abuse, 402-10. Washington, DC: U.S. Government Printing Office.
- McShane, L. 1992. How fried-egg ads fought the drug war. *San Francisco Chronicle*, 19 June, B3, B5.
- Mezvinisky, N. 1961. Scientific temperance instruction in the schools. *History of Education Quarterly* 1:48-56.
- Miles, M. B., and A. M. Huberman. 1994. *Qualitative data analysis*. Newbury Park, CA: Sage.
- Milgram, G. G. 1976. A historical review of alcohol education: Research and comments. *Journal of Alcohol and Drug Education* 21:1-16.
- Miller, M. 1975. *Drug consumer safety program*. Eugene: University of Oregon Drug Information Center.

- Miller, M. and S. Burbank. 1997. *Using alcohol responsibly*. Mosier, OR: Mothers Against Misuse and Abuse.
- Miller, R. L. 1996. *Drug warriors and their prey: From police power to police state*. Westport, CT: Praeger.
- Moore, D. and B. Saunders. 1991. Youthful drug use and the prevention of problems. *International Journal of Drug Policy* 2:29-33.
- Morgan, H. W. 1981. *Drugs in America: A social history, 1800-1980*. Syracuse, NY: Syracuse University Press.
- Moskowitz, J. M. 1989. The primary prevention of alcohol problems: A critical review of the research literature. *Journal of Studies on Alcohol* 50:54-88.
- National Commission on Marihuana and Drug Abuse. 1973. *Drug abuse in America: Problem in perspective*. Washington, DC: U.S. Government Printing Office.
- National Institute on Drug Abuse. 1975. *Doing drug education: The role of the school teacher*. Rockville, MD: National Institute on Drug Abuse.
- Nattress, W. 1893. *Public school physiology and temperance*. 43d ed. Toronto, Canada: William Briggs. The W. J. Gage Company.
- Nellis, M., ed. 1972. *Resource book for drug abuse education*. 2d ed. Rockville, MD: National Clearinghouse for Drug Abuse Information.
- Ormond, C. 1929. *Temperance education in the public schools*. Westerville, OH: American Issue.
- Overton, F. 1897. *Applied physiology—Advanced grade: Including the effects of alcohol and narcotics*. New York: American Book Company.
- Palmer, B. R. 1937. *A syllabus in alcohol education*. 5th ed. Evanston, IL: National WCTU Publishing House.
- Payne, E. G. 1931. *The menace of narcotic drugs*. New York: Prentice-Hall.
- President's Advisory Commission on Narcotic and Drug Abuse. 1963. *Final report*. Washington, DC: U.S. Government Printing Office.
- Rathbone, J. L. 1952. *Tobacco, alcohol and narcotics*. New York: Oxford Life Guidance Pamphlets.
- Reeves, J. L. and R. Campbell. 1994. *Cracked coverage*. Durham, NC: Duke University Press.
- Richards, L. G. 1972. Evaluation in drug education: Methods and results. In *Resource book for drug abuse education*. 2d ed., edited by M. Nellis, 87-90. Rockville, MD: National Clearinghouse for Drug Abuse Information.
- Rorabaugh, W. J. 1979. *The alcoholic republic: An American tradition*. New York: Oxford University Press.
- Rosenbaum, M. 1996. *Kids, drugs, and drug education: A harm reduction approach*. San Francisco: The National Council on Crime and Delinquency.
- Schaps, E., R. DiBartolo, J. Moskowitz, C. S. Palley, and S. Churgin. 1981. A review of 127 drug abuse prevention program evaluations. *Journal of Drug Issues* 11:17-43.
- Skager, R. and J. H. Brown. Forthcoming. Toward a reformation of drug education. In *Drug policy: Psychological, philosophical and legal issues*, edited by J. Fish.
- Stickgold, A. and A. Brovar. 1978. Undesirable sequelae of drug abuse education. *Contemporary Drug Problems* 7 (1): 99-115.
- Stuart, R. B. 1974. Teaching facts about drugs: Pushing or preventing? *Journal of Educational Psychology* 66 (2): 189-201.
- Sullum, J. 1996. Smoke alarm. *Reason Magazine*, May.
- Swisher, J. D. 1979. Prevention issues. In *Handbook on drug abuse*, edited by R. L. DuPont, A. Goldstein, and J. O'Donnell. Washington, DC: National Institute on Drug Abuse.

- Teaching topics on narcotics. 1931. *The Journal of the National Education Association*, February, 61-64.
- Tobler, N. S. 1986. Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. *Journal of Drug Issues* 16:537-67.
- U.S. Department of Education. 1988. *Drug prevention curricula: A guide to selection and implementation*. Washington, DC: Office of Educational Research and Improvement.
- U.S. General Accounting Office. 1991. *Drug abuse prevention: Federal efforts to identify exemplary programs need stronger design*. Washington, DC: U.S. Government Accounting Office.
- Wallack, L. 1980. Mass media and drinking, smoking, and drug taking. *Contemporary Drug Problems*, Spring, 49-83.
- Wallack, L. and K. Corbett. 1990. Illicit drug, tobacco, and alcohol use among youths: Trends and promising approaches in prevention. In *Youth and drugs: Society's mixed messages*. OSAP Prevention Monograph 6, edited by H. Resnik. DHHS Publication No. ADM.90-1689. Washington, DC: U.S. Government Printing Office.
- Watson, M. 1991. Harm reduction—Why do it? *International Journal of Drug Policy* 2 (5): 13-15
- White House Conference for a Drug Free America. 1988. *Final report*. Washington, DC: U.S. Government Printing Office.
- Willard, F. E. 1886. *Woman and temperance or the work and workers of the Woman's Christian Temperance Union*. 5th ed. Chicago: Woman's Temperance Publication Association.
- Worden, M. 1979. Popular and unpopular prevention. *Journal of Drug Issues*, Summer, 425-31.
- Zinberg, N. E. 1984. *Drug, set and setting*. New Haven, CT: Yale University Press.
- Zinberg, N. E. and J. A. Robertson. 1972. *Drugs and the public*. New York: Simon & Schuster.

Many reports of successful school-based intervention programs can be criticized for their choice of a unit of analysis and for the neglect of measurement errors. This article is an illustration of how different conclusions can be reached from different choices of units of analyses and/or of different treatment of the data. This is done by a reanalysis of a well-reported data set. The data is thoroughly taken apart, using different statistical techniques. The result of the analyses shows that earlier reported effects of a normative school-based drug prevention program were not found. The subsequent search for moderator effects of the same program, such as a lowering effect on the relationship between the pre- and posttest or on the relationship between respondents' use and the use of their friends, was not successful either. It is concluded that the null hypothesis of zero effects should be retained. More successful was a search for individual characteristics that show significant relationships with respondents' alcohol use. Among them was the abuse of alcohol by adults in respondents' direct social environment and the use of friends.

AN ILLUSTRATION OF ITEM HOMOGENEITY SCALING AND MULTILEVEL ANALYSIS TECHNIQUES IN THE EVALUATION OF DRUG PREVENTION PROGRAMS

ITA G. G. KREFT

California State University, Los Angeles

INTRODUCTION

School-based drug prevention programs have been a part of the U.S. "war on drugs" campaign during the past 20 years. The most widely used program among them is D.A.R.E.¹ Studies that evaluate the effects of D.A.R.E. are disappointing. Consequently, alternative programs have been developed that try to avoid some real or imagined flaws of D.A.R.E.² My article has two distinct objectives. The first is the evaluation of drug prevention programs,

AUTHOR'S NOTE: *This research was supported by NIDA grant #DA09649-02. Nida grant # DA09649-02 provided me with a total of 6 months during a period of 2 years (1994 -1996) to reanalyze the Adolescent Alcohol Prevention Trial data, collected by Hansen and Graham. The data are made available by John W. Graham. Thanks to W. B. Hansen, and J. W. Graham for making their data available to me.*

EVALUATION REVIEW, Vol. 22 No. 1, February 1998 46-77
© 1998 Sage Publications, Inc.

and the second is the introduction to fairly recent developed techniques, multilevel and homogeneity analysis.

In a review of the literature, I found studies that went beyond the general question: "Do school-based drug prevention programs work or not?" The studies can be divided into two groups: qualitative, where theory is grounded in empirical data, and quantitative, where a theory is applied to make changes in existing programs. One qualitative approach is discussed in the article of D'Emidio-Caston and Brown (this issue). Examples of quantitative approaches, where theory guides the data collection, are found in Graham, Marks, and Hansen (1991; see also Hansen 1993; May 1993), where the components of resistance training programs (which is a D.A.R.E. concept) are defined and compared with social theories (e.g., Bandura 1977, 1986; Jessor and Jessor 1977). Hensen et al. 1991 analyses show that resistance training significantly improved adolescent refusal skills, but the same skills failed to predict less alcohol use. They propose a new program, based on social theory, called "normative education" (NORM),⁴ which seems to be successful in lowering the onset of alcohol use in teenagers (Hansen and Graham 1991).

The conclusion that resistance training alone does not work is reported in other literature as well. An overview by Dukes, Ullman, and Stein (1996) concludes that the effectiveness of D.A.R.E. is inconsistent, where most studies indicate very little or no effects. Ennett et al. (1994) reach the same conclusion. In their meta-analysis of short-term effects of D.A.R.E., they find slight and, except for tobacco use, not statistically significant effects (last cited p. 1398).

The inconsistencies in the reported success of drug prevention programs may be partly due to the methods used to analyze the data. Data are sometimes analyzed at an aggregated class level (see Dukes, Ullman, and Stein 1995; Hansen and Graham 1991), whereas the results are presented as being valid for individual students, resulting in ecological fallacies (see on this topic Robinson 1950; Kreft and De Leeuw 1988, among others). In some analyses, the response variable is dichotomized, where alcohol use is coded as one and non-use as zero (e.g., Hansen and Graham 1991; Graham et al. 1991). In that drug prevention programs deal largely with teenagers, who are experimenting with life, a single experiment will put them in the users/abusers category in this type of coding. Such handling of the data may hide more than it reveals.

Statistical testing is another issue that may lead to rare and nonreproducible findings. Evaluation of a drug prevention program may be based on a single statistical test of significance. There exists a vast literature about the merits of statistical testing, and the interpretation and misinterpretation of such tests (e.g., Wang 1993). A mistake easily made is to think that a statistical

significance between the control group and the program is an indication of large numbers of students saved. But a small p value or a large t statistic is only a confidence statement regarding the rejection of the null hypothesis. If drug prevention evaluations are based on many observations, as they mostly are, a small difference in numbers of students abstaining from alcohol between the program and the control group can result in a significant statistical difference.

Related to the issue of statistical significance is practical significance. The most important information, which is the magnitude of the significant effect, is missing in many of the reports of successful programs. The real difference in numbers of alcohol users in a successful program, as compared to other conditions, is hidden behind F - or t -tests. In the analysis of abstainers in this article, I show that the successful normative program has 9 more abstainers, as compared to the control group. Another example is the reported 62.4% reduction in the rate of onset of drunkenness attributed to the normative program by Hansen and Graham (1991, 422-23). A close look at the cross-break table of the same data shows that this large percentage is caused by only 32 students, 1 out of more than 2,000 students in the study, in that drunkenness is a rare phenomenon at that age (seventh and eighth graders). The cross-break table shows a significant chi-square of 0.04.

Besides, significant results are reported, without mentioning the number of nonsignificant results. Based on statistical theory, we know that 1 out of 20 tests will show significant effects by chance alone. The few effects reported in the literature may suffer from this capitalization on chance, because in this type of research a choice among many response variable is possible. In the data analyzed in this article, I have a choice between 12 variables that measure drinking behavior or attitudes toward drinking. All of these are potential response variables.

The 12 response variables present in the data of Hansen and Graham (1991) are used in this article to create a single summarized response variable, labeled alcohol involvement of respondents. By using a scaling method for categorical data, a continuous scale is obtained, where high use is scaled low (negative) and no use is scaled high (positive). In all the different analyses, this student level variable is used as the response variable.

The report of my re-analysis of the Hansen and Graham data consists of two distinctive parts with two purposes. In the first part of the analysis, techniques are introduced with the purpose of elucidating the statistical problems and fallacies mentioned earlier. Two suitable techniques became available recently in user-friendly software packages. One is especially developed for analyzing hierarchically nested data and allows one to analyze the data at class and student level (ML3 1989). The other technique is useful

for data reduction and the transformation of ordinal-scaled items into a single continuous variable (HOMALS 1989). Using this last technique generally results in a scale with less measurement error compared to the original items.

In the second part, the actual analyses are executed where the effects of drug prevention programs are evaluated. The case for zero effects of programs is made as strong as possible, by using different analysis methods, a technique known as triangulation. Triangulation can show if weak effects are merely happening by chance alone. If an effect is weak, it may show up in one analysis but not in another, which is an indication of a questionable result. All of this ties in with the earlier discussion: If an effect is statistically significant, how large and real is that effect, as expressed in interpretable numbers?

DESCRIPTION OF THE DATA

The data are collected by the Adolescent Alcohol Prevention Trial (AAPT), a longitudinal drug prevention trial examining two psychologically based strategies for preventing the onset of adolescent alcohol and drug use, a longitudinal study that measures students over several years. For this article, the pretest measurement of 1987 in seventh grade of junior high school is used and the measurement is used 1 year later, after the implementation of the treatment, in eighth grade. The 12 schools in the study consist of 118 school classes. Seven of the 12 schools are public schools, whereas 5 are Catholic schools. Schools are randomly assigned to one of four experimental conditions, which are

1. A control group of 32 classes receiving the general information program about consequences of use only (CONTROL)
2. A group of 33 classes receiving resistance training (RT^3) + information
3. Normative education (NORM) + information was received by 27 classes
4. A group of 26 classes receiving a combined condition, with RT + NORM + information (BOTH) (see Hansen and Graham 1991, for more detailed information)

The pretest sample consisted of 3,027 students, and the posttest data contains the answers of 3,147 respondents; 2,378 students answered both questionnaires, one in 1987 and one in 1988. For analysis purposes, two levels are recognized, the students as the first level and classes as the second level of observation. Classes are as equally important as a level of analysis as students are for this study. A class climate can be one of the influences that

make drug prevention programs a success or a failure, and for that reason should not be ignored in the evaluation of such programs. Information at both levels is present in the data. The most important class-level measurement is the drug prevention program, which has four different categories, as described above. The programs are dummy coded, with the control group as reference. At student level, more than 200 items are present. For the analyses in this article, 33 are picked to construct four new variables. The four variables are

- pretest for alcohol involvement of the respondent,
- the respondents' posttest (which is the response variable),
- the alcohol use of friends, and
- the alcohol use by adults in the respondents' environment.

Several of these student variables are averaged and used as an indication of class climate. The most important goal of the analyses is to find effects of drug prevention programs, and especially of the reported successful program NORM. It is known that programs can have general effects on students' behavior, but also moderator effects. Moderator variables are defined in regression analyses as interacting variables, interacting with the relationship between explanatory and response variables. To investigate if moderator effects are present in these data, a new type of variable is created, known in the multilevel literature as crosslevel interactions. The name "crosslevel" indicates that the interaction involves variables from different levels of the hierarchy, here the student level and the class level. By multiplying two variables, one from each level, crosslevel interaction terms are constructed for the purpose of investigation. The program NORM is used to create two interactions by multiplying this variable with two student-level variables, the respondents' pretest and friends' alcohol use. If a significant moderator effect of NORM is present in the data, it will be an indication that this program has an effect on the relationship of the interacting variable and the response variable. When that effect is significant and negative, the conclusion can be drawn that NORM has a moderator effect by lowering the relationship between pre- and posttest, or between friends' alcohol consumption and respondents' posttest alcohol involvement.

The surveys are collected 1 year apart, first in seventh grade, and 1 year later, when the same students are in eighth grade. In both years, the same questionnaire is administered, containing many items that measure alcohol use of respondent, use by friends, and use by adults in the social environment of the respondents. The two data sets combined (the one measured in 1987

and the one in 1988) resulted in a data set with a total of 2,378. Due to missing cases in one or the other year, 649 students are dropped from the analysis.

DESCRIPTION OF ANALYSES TECHNIQUES

The two techniques used to analyze these data are briefly introduced in the next paragraphs, because it is expected that they are relatively unknown to most readers. For a more extensive description, as well as an explanation of the new modeling concepts behind these techniques, I refer to the relevant literature. For details of applications to drug prevention research, I refer to Kreft (1994, 1997).

MULTILEVEL ANALYSIS

Multilevel analysis is especially developed for data collected in situations where observations are clustered in groups. In the evaluation of drug prevention programs, observations are mostly clustered within programs, whereas programs are administered to existing groups, such as school classes, health groups, or community centers. The goal of drug prevention programs is to influence individual behavior, although randomization and implementation of the intervention occur at the group level. In our data, the individuals are junior high school students in seventh and eighth grade, clustered in 118 school classes. Prevention programs are applied at the class level. Students and classes are both important levels of observation. First, existing classes are not equal to randomized groups but have over time developed their own class climate and class dynamics. As a result of this interaction among classmates, students in the same class are more alike than students in different classes, resulting in dependent observations. Also, students in the same class share behaviors and value systems that may interfere with the failure or success of prevention programs. In multilevel analyses, both levels are defined as levels of influence and both exert influence on the behavior of students. In these data, even more levels are present; the first level is the individual student, nested within the second level, the class, whereas the class is nested in a third level, the school. For the analyses in this article, only two of the three levels are considered, the student and the class. The class level is chosen over the school level because of its importance in relation to drug prevention programs. The school level could be included as a third level, but because only 12 schools are present in these data, the school level is ignored.

The two levels analyzed in the multilevel analyses in this article are students (Level 1) and classes (Level 2).

Before any analysis techniques were available to analyze both levels at the same time, discussions centered around the question, What is the appropriate level of statistical analysis for these type of data, student level or class level? (Robinson 1950; Hannan 1974; Burstein 1980; Kreft and De Leeuw 1988). The conclusion reached in the literature of that time is that class-level analyses can deliver the wrong message in relation to student behavior, in that it can only make conclusions about classes. Inferences made about individuals based on aggregated data can lead to incorrect conclusions, as illustrated by Robinson (1950), who labeled this the "ecological" fallacy. Robinson's examples show that aggregated correlations can have an opposite sign compared to individual correlations, even when calculated over the same data. For the same effect on regression coefficients, see Kreft and De Leeuw (1988), who labeled this the "see-saw" effect.

Multilevel analysis does take that intraclass correlation into account, and analyses can be done at student level while also analyzing the data at class level. One of its main advantages is that many more research questions can be answered, as compared to single-level analysis. It allows to explore the data at both levels and to discover the complex relations of variables measured at different levels with the alcohol use of respondents. Research questions that mix levels of influence together are common in multilevel analyses. For example, the question, "What is the effect of class climate, together with the effect of alcohol consumption of friends, on respondent's alcohol use?" can be answered. Aggregation of student-level measurements to the class level deletes the important variation in respondents' personal relationships with family and friends. In research where important variation in social environment of respondents is deleted, "at-risk" students can no longer be identified. My conclusion is that student-level variation cannot be ignored, but neither can class variation. Drug prevention programs are most often administered to whole school classes, where a certain class climate may interfere with the success of any of these programs. In the multilevel analyses reported later, both levels are taken into account. Questions related to students' individual behavior, as well as the behavior of school classes, are asked and answered.

HISTORY OF MULTILEVEL MODELS

Multilevel regression models have been developed for analysis of hierarchically nested data, such as students nested in classes. These techniques

correct for intraclass correlation and allow the researcher to estimate individual effects, as well as group-level and drug prevention program effects. In the models, it is possible to test cross-level interactions, such as between drug prevention program conditions and student characteristics. Such effects are also known as moderator effects, as discussed earlier.

Multilevel methods are presently applied in many different fields, such as education (Aitkin and Longford 1986; De Leeuw and Kreft 1986; Raudenbush and Bryk 1986), public health (Laird and Ware 1982; Longford 1987), and psychiatry (Hedeker, Gibbons, and Davis 1991; Hedeker, Gibbons, and Flay forthcoming). At the moment, many programs are available for this type of analysis (see Kreft, De Leeuw, and van der Leeden 1994, for a review of some of these programs). Carbonari et al. (1994, 89) consider multilevel methods "an important advance in the field." They specifically mention as a virtue of the new model the resolution of the issue of the unit of analysis for unbalanced data.

For this article, the software ML3 (1989) is used.

SCALING VARIABLES WITH HOMOGENEITY ANALYSIS

The second technique described is homogeneity analysis, a data reduction technique that scales several variables into a single continuous variable. Four scales in total will be constructed, where two of the most important ones are respondents' alcohol involvement in pre- and posttest. Both are measured with 12 variables. In 1987, and again in 1988, 12 questions are asked related to present and anticipated alcohol use. All 12 are used to construct one pre- and one posttest measurement, representing respondents' alcohol involvement in both years. The following are the 12 questions (for more details, answer categories, and frequencies, see the Appendix):

- Item 19: How many drinks of alcohol have you had in your whole life? 1 = *none*, 9 = ≥ 100
- Item 20: How many drinks in the past month? 1 = *none*, 8 = ≥ 20
- Item 23: How many drinks in the past week? 1 = *none*, 8 = *11 or more*
- Item 24: How many days in the past month did you have a drink? 1 = *none*, 6 = *15 to 30*
- Item 25: How long since you had a drink of alcohol? 1 = *less than 24 hours*, 7 = *never a drink*
- Item 28: Think of the day during the past month when you drank the most alcohol. How many drinks did you have on that day? 1 = *I never drink*, 8 = *5 or more*

TABLE 1: A Comparison of Two Items Measuring Alcohol Use of Respondents

Categories Item 20	Categories Item 24					
	None	1	2 or 3	4 to 7	8 to 14	15 to 30
1. None	2,347	28	1	4		
2. Only a sip (for religious purposes)	116	8	6			
3. Only sips	89	104	31	5	1	1
4. Part of a drink	17	56	23	3		
5. 2 to 4	13	36	44	16	3	3
6. 5 to 10	4	4	6	15	5	1
7. 1s1 to 20		2	3	3	4	1
8. More than 100		1	3	2	1	6

NOTE: The inconsistent answers are in boldface type. Item 24: How many days in the past month have you had alcohol to drink? Item 20: How many drinks of alcohol have you had in the past month?

- Item 29: How often do you imagine having a drink? 1 = *often*, 4 = *never*
- Item 32: Do you think you will drink in the next months? 1 = *yes*, 4 = *no*
- Item 33: Do you think you will ever drink alcohol every day? 1 = *yes*, 4 = *no*
- Item 34: Do you think you will ever drink every month? 1 = *no*, 4 = *yes*
- Item 35: How many times have you ever been drunk? 1 = *never*, 6 = *≥ than 20 times*
- Item 38: Do you think you will get drunk in the next month? 1 = *no*, 2 = *yes*

Homogeneity analysis is a method developed for the analyses of categorical variables, as most variables in the list above are. The technique is used in my example to construct a numerical scale based on all items present. The software used is HOMALS, which is an acronym for homogeneity analysis by alternating least squares, and available in Categories (SPSS, 1989). This type of analysis is comparable to principal component analysis (PCA), but for categorical instead of numerical data. The method has been given many names, because it was discovered independently by different people but first applied by Guttman (1941) for the scaling of constructs. The best known name for this technique is multiple correspondence analysis, used by Benzecri (1973) and Greenacre (1984). Other names are dual scaling, method of reciprocal averaging, linearization of regression, and seriation (see Van de Geer 1993a, 1993b). If all variables are binary, results of HOMALS will be the same as those obtained from classical PCA.

The technique is first demonstrated using the variable for alcohol involvement of respondents, in pretest and in the posttest year. As illustrated in

Table 1, items measuring the same behavior are not always in agreement. The table shows that two items, measuring the same behavior, contain measurement error. Some of the respondents give answers that do not agree with previous answers given.

Many inconsistencies are found in Table 1, where 33 students (see the boldface numbers 28, 1, and 4 in the first row) report not drinking last month, yet answer for Item 24 that they had a drink at one or more days of that month. The reverse is also present. Of the students that answer to Item 20 that they did not drink that month, except maybe a sip, 85 students answer to Item 24 that they had a drink at several occasions that month, where answers range from 1 to 15 and more (see the boldface numbers in the first three rows, 28, 1, 4, 8, 6, 31, 5, 1, 1). More inconsistencies are present, all illustrating that items measuring the same behavior are not always answered in the same way. The HOMALS construct for "alcohol involvement" of respondents is based on the total answering pattern over all items, which takes inconsistencies in answering patterns into account. As a result, a more reliable scale for alcohol involvement becomes available.

Scaling the variables with homogeneity analysis is also useful for other purposes. The technique can deal with missing data, preventing listwise deletion of cases when respondents have one or more answers missing. The missing data are replaced by values that are close to the values for students with similar but complete answering patterns.

ANALYSES RESULTS OF SCALING

The items that measure pre- and posttest alcohol involvement are skewed to the right with a mode at Category 1, no alcohol use. This skewness is present in most data on drug use in such a young population. The proportion of abstaining students on the items is between 55% and 95% (see Appendix), depending on which of the 12 questions in the pre- or the posttest is observed. The lowest number of abstainers is found for the item that measures alcohol use over a lifetime (Item 19), followed by the percentage of abstainers for drinking in the past months (Item 20). The highest percentage of abstainers is found in the item measuring alcohol use in past week (Item 23). Combining the 12 questions yields a single variable with a smaller proportion of abstainers than most of the separate items have, also resulting in a variable with better statistical properties.

TABLE 2: Eigenvalues and Discrimination Measures for the 1987 and 1988 Smoking Variables

	<i>Discrimination Measures^a</i>		<i>Eigenvalue^b</i>		<i>Number of Observations</i>	
	1987	1988	1987	1988	1987	1988
Item 19	0.670	0.712	0.5566	0.6081	N = 3,027	N = 3,047
Item 20	0.709	0.775				
Item 23	0.550	0.580				
Item 24	0.694	0.737				
Item 25	0.660	0.694				
Item 28	0.658	0.715				
Item 35	0.516	0.562				
Item 38	0.460	0.586				
Item 29	0.421	0.498				
Item 32	0.654	0.678				
Item 33	0.232	0.249				
Item 34	0.456	0.511				

a. The correlation of each variable with the underlying scale.

b. A measure for the reliability of the scale.

RESULTS OF HOMOGENEITY ANALYSES FOR SCALING OF RESPONDENTS' ALCOHOL PRE- AND POSTTEST

In the first two analyses, reported in Table 2, the scales are constructed for the pre- and posttest, and labeled respondents' alcohol involvement. A conceptual difference among the items is indicated in the table by a line that divides the first eight items from the last four items. The first items measure actual alcohol consumption, whereas the last measure drinking as projected in the future.

The table shows discrimination measures and eigenvalues. Discrimination measures are the factor loadings for each variable for each year, labeled discrimination measures. The eigenvalue for an analysis is a measure of overall fit, one for each year. Although the concept of discrimination measures is equal to the concept of factor loadings in traditional PCA, it would be misleading to use the same name, in that the estimation methods are not comparable among methods. The discrimination measures in Table 2 show different values or loadings. These different values indicate the different contributions of variables to the underlying scale. Item 20 has the highest discrimination measure in both analyses, showing that this question (alcohol consumption during the last month) is the most important one.

The magnitude of the discrimination measures shows if an item is an important contributor to the scale formed by all variables together. The item about weekly drinking (Item 23) and about being drunk (Item 35) contribute somewhat less to the scale, as indicated by the magnitude of the discrimination measure. The reason these items do not discriminate equally strong as the other items is due to the high number of abstainers. The same is true for Item 33 that measures the most extreme drinking behavior, "Do you think you will ever drink alcohol every day?" This question is answered by most students with "no" (see the Appendix), resulting in a low discrimination measure. Items may not discriminate very highly, but they are not deleted from the analysis. This decision is based on theoretical considerations, in that all three items are strong in measuring alcohol involvement.

Discrimination measures can also be interpreted as the correlation with the underlying scale. In the two separate analyses in Table 2, Item 20 has the highest correlation with the newly formed scale in 1987 as well as in 1988. Item 33, on the other hand, has the lowest correlation in both years. The two analyses show similar patterns in the other variables as well, an indication of the reliability of the scales. A discrimination measure of a variable shows the proportion of variance of the variable that is between categories of that variable. Or, equivalently, a discrimination measure is the variance between the category quantifications of that variable. Consequently, a low value shows that the categories of that variable do not discriminate much.

The items that are the strongest determiners of present and future alcohol use are the items that measure monthly drinking (Items 20, 24, 28, and 32), together with lifetime drinking (Item 19). These items are also theoretically of the most importance, which makes the new constructed variable for pre- as well as for posttest a valid measure for our analyses of the evaluation of drug prevention programs.

Table 2 shows the eigenvalues for both analyses, which are 0.5566 and 0.6081, respectively. Eigenvalues are average discrimination measures and can be used as an overall measure of fit. The highest possible value of eigenvalues (and discrimination measures as well) is 1.00. In the table, the eigenvalue of 0.5566 for the analysis of items measured in 1987 indicates that 56% of the variation of the new scale is between categories of all variables. The items measured in the posttest explain 61% of the variation.

CATEGORY QUANTIFICATIONS

The category quantifications of each item, part of HOMALS output, are reported in the Appendix. It shows the new category values, which have a

reverse ordering compared to the old category quantifications. The new quantifications indicate that high positive for pre- as well as posttest indicates no alcohol use, whereas low negative indicates a high level of alcohol use. The same is true for the resulting scale, where a positive score means no drinking, or positive behavior, and a negative score is alcohol use or abuse, or negative behavior. For analyses purposes, the new category quantifications of the variables in the analyses are of no further use, but they are of interest for a better understanding of the new scale. The new scale has a mean of zero and a standard deviation of one.

A comparison between the HOMALS category quantifications and the original ones shows that categories of an item are no longer equally distanced, the way the original categories are. For instance, Item 23 (How many drinks in the past week) shows that the HOMALS distance between no drinking and "a sip for religious purposes" is smaller and no longer equal to the distance between "one sip" and "one drink." The new category quantifications behaves very similar among the two analyses, an indication of the reliability of this scaling method as applied to this data.

In the next paragraphs, variables are constructed that measure the alcohol use in the immediate environment of the respondents. One is friends' alcohol use; and the second is a construct for alcohol use by adults in the immediate environment of the respondents. HOMALS is used again to summarize the available items.

CONSTRUCTION OF A PRETEST FOR FRIENDS' ALCOHOL USE

In the same way as before, all variables that measure alcohol involvement of friends of the respondents are used to form a scale. Seven questions are available in the data about friends and their alcohol consumption and behavior in the pretest 1987. These questions are

- Item 1: How many of your three best friends have ever tried drinking alcohol?
- Item 2: How many of your best friends have had alcohol to drink in the past month?
- Item 2a: Have your best friends in your grade in this school had alcohol to drink in the past month?
- Item 3: How many of your best friends have ever been drunk?
- Item 4: How many of your three best friends have been drunk during the past month?
- Item 36: How often are you with kids who are drunk?

TABLE 3: Eigenvalues and Discrimination Measures for 1987 and 1988 for Friends Drinking Alcohol

	<i>Discrimination Measures^a</i>	<i>Eigenvalue^b</i>	<i>Number of Observations</i>
Item 1	0.572	0.582	N = 3,027
Item 2	0.731		
Item 2a	0.565		
Item 3	0.634		
Item 4	0.570		
Item 36	0.420		

a. The correlation of each variable with the underlying scale (similar to a component loading).

b. A measure for the reliability for homogeneity of the scale.

In Table 3, the results are reported of the homogeneity analyses with the six available items measuring alcohol use of friends. The discrimination measures of the separate items show that Items 2 and 3 are the most important contributors. They measure receptively drinking of friends in general and during the past month. The scale formed by the items is scaled in the same way as the pre- and the posttest, with a mean of zero and a standard deviation of one. The new variable is labeled "friends alc" in the analyses reported later in this article.

CONSTRUCTION OF A SCALE FOR ALCOHOL USE IN THE IMMEDIATE SOCIAL ENVIRONMENT

The next scaling is a construct based on three variables that measure drinking by adults in the environment of the respondent. For the construction of this scale, three questions from the 1987 questionnaire are used:

- Item 26: How many times have you been offered a drink of alcohol in the past month?
- Item 30: How often are you with adults who are drinking alcohol?
- Item 37: How often are you with adults who are drunk?

The scale formed by the three variables is again scaled with high positive indicating low alcohol use, whereas high negative indicates high alcohol use. The name used for this variable in the analyses is "social."

In Table 4, the results of this homogeneity analysis are reported, which show that the middle question, Item 30, "How often are you with adults who

TABLE 4: Eigenvalues and Discrimination Measures for 1987 for Drinking of Alcohol in the Environment

	<i>Discrimination Measures^a</i>	<i>Eigenvalue^b</i>	<i>Number of Observations</i>
Item 26	0.641	0.5457	N = 3,027
Item 30	0.312		
Item 37	0.684		

a. The correlation of each variable with the underlying scale.

b. A measure for the reliability of the scale.

are drinking alcohol?" contributes the least to the scale. This supports the finding of Brown and D'Emidio-Caston (1995) that there exists a difference between use and abuse. Drinking of alcohol can happen in a social context and does not have to be abusive. Alcohol abuse is indicated more in the questions stated in Items 37 and 26, which contribute equally strong to the scale, with discrimination measures of 0.641 and 0.684, respectively.

TRIANGULATION BY USING DIFFERENT TECHNIQUES

TRADITIONAL ANALYSES: MULTIPLE REGRESSION AND ANALYSIS OF COVARIANCE

In the introduction, I argued that analyzing these clustered data at one single level is a mistake, in that one or the other level will be ignored. On the other hand, I expect that strong effects will even show up in "flawed" methods, whereas weak effects may not always survive in different analysis techniques. This method is known as triangulation. A comparison of effects based on different ways to analyze the data can support an earlier found effect, or it fails to support the earlier findings. If that happens, the earlier findings are made questionable.

The first analysis technique presented in this paragraph is multiple regression with students as the unit of analysis. The second technique is analysis of covariance (ANCOVA), where the programs are the factors, students' alcohol involvement in 1988 is the response variable, and the alcohol involvement in 1977 is the covariate. The results obtained by these two traditional linear techniques are compared with the results obtained by a multilevel analysis.

All models in this section are simple. In the first model (Model 1), the pretest predicts the posttest. In a next model (Model 2), the three drug

TABLE 5: Multiple Regression Between Pretest Homals-Alcohol and Posttest Homals-Alcohol and Two Explanatory Variables, Prevention Programs RT and NORM, $N = 2,378$

Response Variable	Alcohol88 (Homals)					
	Model 1			Model 2		
	Coefficient	SE	R ²	Coefficient	SE	R ²
Constant	-0.03	0.02		-0.03	0.02	
Alcohol87 (Homals)	0.65	0.02**	0.38	0.65	0.02**	0.38
NORM				0.11	0.05*	
RT				-0.05	0.04	
BOTH				0.02	0.04	

* $p < .05$. ** $p < .01$.

prevention programs are added. The interpretation of the analysis results follow the same scenario. First I look at the value of the individual coefficients in the model and compare these with their standard errors. When an individual coefficient is significant, it is indicated with an asterisk. But more important, the total model fit is compared among models. Decisions about significant effects are made based on the total fit of a model, rather than on significant tests of individual coefficients. Decisions based on individual coefficients may increase the chance of Type I errors.

Multiple Regression

The first analyses are in Table 5, where a traditional multiple regression analysis is executed, using the two models described earlier, and using the new constructed variables (HOMALS).

In the multiple regression analysis, a statistically significant effect for NORM is present. This result is not supported by a better model fit, because the R^2 for both models (with and without programs) is equal. The explained variance in both regression models is $R^2 = 0.38$. Individual parameter interpretation would be capitalizing on chance, because the overall fit of the model does not show improvement.

Regression analysis is not the best model for the analysis of clustered data. Analysis of covariance, which treats the program conditions at the correct level, is a better approach. It still does not correct for intraclass correlation, but it can serve as a preliminary check for group effects before executing a multilevel analysis.

TABLE 6: Model 1: ANCOVA With Pre- and Posttest Alcohol

	SS	F1	Significance of F1
Pretest	876.13	1456.43	0.000
Program	7.54	4.18	0.006
<hr/>			
<i>Program Conditions</i>	<i>Model 1</i>		
Mean control group	-0.01	N = 671	
Mean RT	-0.06	N = 654	
Mean NORM	0.10	N = 462	
Mean BOTH	0.01	N = 591	

NOTE: SS1 = sum of squares; F1 = F-statistic.

ANALYSIS OF COVARIANCE

In analogy to the regression models in Table 5, an analysis of covariance is executed. In Table 6, results are reported of a model with the posttest (1988) as the response variable and the pretest (1987) as covariate. The factor Program has four conditions, Control, RT, NORM, and BOTH.

The results in Table 6 show that the factor Program, with its four categories, has a significant effect ($p = 0.006$). The bottom half of the table shows the adjusted means for each of the conditions. Remember that the pre- and posttests are constructed so that a high positive score is low or no alcohol involvement, and high negative is high alcohol involvement, with a mean of zero and a standard deviation of one. The four means show that the overall significant effect for the factor Program is partly due to the negative effect of the drug prevention program RT. RT has the largest deviation (-0.06) from the overall mean (0.00), making it the program with the highest mean alcohol use. The control group has a mean of around zero (-0.01). The conclusion based on this analysis can be different from the one obtained with multiple regression. In the regression analyses, the model fit was not improved by adding the program conditions, whereas in ANCOVA, the F test shows a significant effect for the factor Program.

The significant effect of Program is mainly due to the large difference between RT and NORM (a difference of 0.16), less than the difference between NORM and the control group (a difference of 0.9).

Because it is still unclear what to think of the program NORM, the percentage of abstainers for the four program conditions are calculated and reported in Table 7. Abstainers are defined as the percentage of students that

TABLE 7: Percentage of Abstainers for Item 19 in 1988

	<i>Control</i>	<i>BOTH</i>	<i>RT</i>	<i>NORM</i>
Item 19	74%	76%	72%	77%
	(361 out of 489)	(342 out of 449)	(323 out of 447)	(248 out of 324)

NOTE: The percentages are not abstainers over the total sample, but the percentage abstainers of the original group of abstainers a year earlier. Hence, smaller numbers are reported for each group than in Table 6.

do not drink in 1987 and still do not drink a year later, in 1988. If fewer students start drinking in one of the program conditions, as compared to the control group, that drug prevention program is successful in refraining more students from drinking. The percentages are calculated based on Item 19: "How many drinks did you have in your life?" This question is answered in 1987, as well as in 1988. Students who reported that they had a sip to drink are counted as abstainers.

The percentages reported in Table 7 show that for all programs including the control group, the number of abstainers is less in 1988 compared to the previous year. If the percentage abstainers in the RT, NORM, and BOTH is compared with that of the control group (74%), it shows again that RT is the least successful condition (72%), whereas NORM is the most successful (77%). The difference between the RT and NORM is statistically significant, but that is an irrelevant conclusion, because programs need to be compared with the control group. The control group shows a difference of 3% with the NORM program, which is a difference equal to 9 students. It is obvious that such a small difference is neither statistically nor practically significant, more so because the comparison is based on one single item, which most likely contains measurement error, as was illustrated in Table 1. The item is also a self-report, and "one must always be cautious when interpreting analyses based on a single method of measurement" (Donaldson, Graham, and Hansen 1994, 212).

MULTILEVEL ANALYSES WITH RESPONDENTS' PRE- AND POSTTEST

Because the data are based on observations nested in existing classes, intraclass correlation may be present. Intraclass correlation affects the standard errors of regression coefficients in a way that leads to an underestimation. As shown in Barcikowski (1981), the presence of an intraclass correla-

TABLE 8: Multilevel Analysis With Alcohol 1988 as the Response Variable

	<i>Model 1</i>		<i>Model 2</i>	
	<i>Par-Estimate</i>	<i>SE</i>	<i>Par-Estimate</i>	<i>SE</i>
Intercept	-0.03	0.02	-0.03	0.03
Pretest	0.68	0.03**	0.68	0.03**
RT			-0.05	0.04
NORM			0.07	0.04
BOTH			-0.004	0.04
<hr/>				
	<i>Par-Estimate</i>	<i>SE</i>	<i>Par-Estimate</i>	<i>SE</i>
Variance level 1	0.56	0.02**	0.62	0.02**
Variance level 2				
Intercept	0.01	0.01	0.02	0.01*
Slope	0.05	0.01**	0.05	0.01**
Covariance	-0.03	0.01*	-0.03	0.01*
Deviance	5458.21		5451.24	

* $p < .05$. ** $p < .01$.

tion has an effect on the alpha level of the F test in analyses of variance, leading to too liberal tests of significance. A small significant effect of drug prevention program NORM is found in the traditional analyses, which makes the expectation that such an effect will show up in a multilevel analysis doubtful, due to stricter tests of significance in a method that takes intraclass correlation into account. Based on these results as well as on the table with abstainers, no main effects of programs are expected to be found in the multilevel analysis. For reasons of comparison, the programs NORM, RT, and BOTH are included as explanatory variables in the reported multilevel analyses.

Because the previous analyses have indicated that it is most likely that main effects of drug prevention programs will not be found, analyses are used to test if moderator effects of program conditions are present in the data, especially of NORM. Several models test the theory that effects of programs are not equal for all students but have interaction effects. In the literature, it is suggested that interactive effects may exist between drug prevention programs and individual characteristics, although "traditional analyses fail to detect important effects" (see Donaldson et al. 1995, 5). Multilevel analyses are suited to test interactions, such as the moderator effect of NORM in lowering the strength between pre- and posttest.

FITTING THE SAME MODEL AGAIN

The traditional analyses show no promising results for drug prevention programs. This is verified with a multilevel analysis using the same variables as in Tables 5 and 6) where pretest is predicting posttest, together with the three programs NORM, RT, and BOTH. Programs are defined in multilevel analyses as second-level explanatory variables, in that they are measured at the class level. The control group is set to zero. Results are reported in Table 8.

Both models in Table 8 have only pre- and posttest as student-level explanatory variables, whereas in the second model (Model 2), the three program conditions are added to the model. The results show that again pretest predicts the posttest, but program conditions do not add significantly to an explanation of the posttest variation. Several sources in the table support that finding. The individual coefficients of the three conditions, RT, NORM, and BOTH, show no significant effects after correction for pretest (see Model 2 in Table 8). A better way for testing the effects is looking at the model fit. Model fit can be checked by taking the difference in deviance between Model 1 and Model 2, which is 6.93, and comparing this difference to the degrees of freedom (*df*) lost. Comparing 6.93 with 3 *df* shows that Model 2 does not significantly improve the fit compared to Model 1. The same conclusion is reached here as in the regression analysis (see Table 5) where the R^2 did not change by adding the three program conditions. There is still a third way to check if adding program conditions improves the model, which is by comparing the variance component of the intercept over models. The variation in the intercept in Model 1, the model without the program conditions, is not significant (0.01, with a standard error of the same magnitude). The zero variance of the intercept indicates that no differences in the mean level of posttest alcohol involvement over school classes is present. In sum, I found that school-level characteristics, including programs, cannot explain school class variation. All classes behave in the same way, after correction for pretest. The only promise in Model 1 is in the significant slope variation of the pretest, 0.05, with a standard error of 0.01. In the following models, I will try to explain this variation among school classes by adding cross-level interactions with one of the program conditions and pretest.

MULTILEVEL ANALYSES WITH RESPONDENTS' PRE- AND POSTTEST, "FRIENDS" AND "SOCIAL"

In the following analyses, the earlier constructed variables for friends' drinking behavior and the behavior in the social environment of the respon-

TABLE 9: Multilevel Analysis With Alcohol 1988 as the Response Variable

	<i>Model 1: Main Effects</i>		<i>Model 2: Cross-Level Interactions</i>	
	<i>Par-Estimate</i>	<i>SD</i>	<i>Par-Estimate</i>	<i>SD</i>
Intercept	-0.03	0.03	-0.04	* 0.03
Alcohol pretest	0.52	0.03**	0.55	0.03**
Friends alc	0.19	0.02**	0.19	0.02**
Social alc	0.06	0.02**	0.06	0.02*
RT	-0.05	0.04	-0.04	0.04
NORM	0.07	0.04	0.11	0.05*
BOTH	-0.02	0.04	-0.02	0.04
Interaction Norm * pretest			-0.10	0.06
Variance level 1	0.53	0.02**	0.53	0.02**
Variance level 2				
Intercept	0.01	0.01	0.01	0.01
Alcohol slope	0.05	0.01**	0.04	0.01**
Covariance	-0.03	0.01*	-0.03	0.01*
Deviance	5,350.65		5,348.09	

NOTE: Model 1 with main program effect. Model 2 with cross-level interaction between NORM and alcohol pretest.

* $p < .05$. ** $p < .01$.

dents is added. To test the hypothesis that the most successful program NORM has a lowering effect on the relationship between pre- and posttest, a cross-level interaction is introduced in the model (see "Interaction Norm*pretest" in Table 9). I expect that the program lowers the strength of the relation between pre- and posttest.

At the student level, the posttest construct for alcohol is used again as response variable. As explanatory variables, pretest alcohol use, friends' alcohol involvement (friends), and the social context (social) measuring adults that are drunk or offer drinks are added. At the school class level, the three drug prevention programs RT, NORM, and BOTH are again included as explanatory variables. In Model 2 of Table 9, the hypothesis is tested that NORM has a cross-level interaction effect with pretest. If a significant interaction is present, lowering the relationship between pre- and posttest, the hypothesis is supported that drug prevention effects are indirect. Model 2 shows an interaction effect that is in the right direction (negative) but it is not significantly different from zero. All student level effects are highly significant.

A suppressor effect is also present in Model 2, where the addition of the interaction term has changed the coefficient of NORM and its standard error in a way that NORM became a significant effect in Model 2. This significant

effect is an artifact of regression models when correlated variables with opposite effects are added to a model. In the example, the interaction coefficient has a sign in the opposite direction of the NORM coefficient, thus enhancing the effect of NORM.

The fit of a model is greatly improved by the addition of the two student level variables, social and friends. Comparing model fits (Model 2 in Table 8 compared with Model 1 in Table 9), an improvement of fit of $5,451.24 - 5,348.09 = 103.15$ is found. Compared to the loss of 2 *df*, this is by all standards a very significant improvement.

The analyses so far have not found strong evidence of effects of drug prevention programs, alcohol involvement is found neither in main effects nor in cross-level interaction effects.

In the next analyses, I test the hypothesis that high mean levels of pretest alcohol involvement in school classes are interacting with prevention program conditions, as was hypothesized by Hansen and Graham (1991) and Graham et al. (1991).

MULTILEVEL ANALYSIS WITH MEANS

Because no main effects of drug prevention programs are expected to be present in these data, the analysis proceeds as an exploration of possible interaction effects of drug prevention programs, interactions with special types of students or with mean levels of alcohol use in school classes. In the literature, it is suggested that such interactive effects exist between drug prevention programs and individual characteristics, although "traditional analyses fail to detect important effects" (see Donaldson, Graham, and Hansen 1994, 5). Multilevel analysis is designed to test this type of "cross-level"⁴ interactions, such as the effect of the most promising drug prevention program NORM on pretest. The hypothesis is that NORM will lower the strength of the relationship between pre- and posttest. If such an effect is found, NORM has a moderator effect, lowering the effect between individual characteristics.

One of the goals of the normative education curriculum (NORM) in the AAPT study is to demonstrate to students that the actual use among students in the school is much lower than students think or perceive. That is, the common statement that "everyone is doing it" is simply wrong. However, if there happens to be relatively more drinking and other substance use in a particular classroom, the credibility of this normative education message may be seriously undermined. Hansen and Graham (1991) wrote: "It has long been suspected that peer pressure is a major cause of onset of use of common

TABLE 10: Multilevel Analysis With Cross-Level Interaction With NORM and Mean

	<i>Model 1: Class Mean of Respondent</i>		<i>Model 2: Class Mean of Friends</i>	
	<i>Par-Estimate</i>	<i>SE</i>	<i>Par-Estimate</i>	<i>SE</i>
Intercept	-0.03	0.03	-0.03	0.03
Alcohol pretest	0.52	0.03**	0.52	0.03**
Social alc	0.06	0.02*	0.06	0.02*
Friends alc	0.19	0.02**	0.19	0.02**
RT	-0.05	0.04	-0.05	0.04
NORM	0.07	0.04	0.07	0.04
BOTH	-0.02	0.04	-0.02	0.04
Interaction Norm*				
Mean-friends alc			-0.06	0.10
Interaction Norm*				
Class Mean Resp	0.14	0.11 (<i>ns</i>)		
Variance level 1	0.53	0.02**	0.53	0.02**
Variance level 2				
Intercept	0.01	0.01	0.01	0.01
Alcohol slope	0.05	0.01**	0.05	0.01**
Covariance	-0.03	0.01	-0.03	0.01
Deviance	5,349.22		5,350.25	

NOTE: The deviance of the same model, but without a cross-level interaction is 5,350.65.

* $p < .05$. ** $p < .01$.

substances" (p. 425). Based on this notion, the hypothesis is tested that school classes with high alcohol use have lower or no program effects, compared to classes with low average alcohol involvement. This hypothesis is tested by constructing two averages for the amount of alcohol involvement in the class, the class means for pretest and the mean for friends' alcohol involvement. These two means are interacting with the program NORM in the next models (see the terms "Interaction Norm" with "Mean-friends alc," in Model 2 and "Interaction Norm" with "Class mean of respondents" in Models 1 and 2 in Table 10).

The hypotheses tested in the Models of Table 10 is that in classes where alcohol involvement is high, program NORM effects are less. The theory underlying the program NORM is that overall high use in a school class interacts with the program effectiveness of NORM. This is measured in the interaction term between NORM and the class "Mean Friends" and "Class Mean Resp" in Table 10. If these interactions are significant and negative, it supports the hypothesis that a higher mean level of alcohol in a class lowers the positive effect of NORM. Because the coefficients for the interaction

terms are not statistically significant, the hypothesis is not supported. The results of both analyses in Table 10 show again that pretest social and friends have highly significant coefficients, whereas program conditions have not. In both models, the interactions between NORM and mean levels of alcohol involvement are not significant. The results in Table 10 do not add any new information to what we already obtained from Table 9. The same is obvious from the model fit compared among models. Using the deviance again, with *df* it shows that Model 1 of Table 9 is very close in deviance to the one observed in Table 10, which are, respectively, 5349.22 and 5350.25. The deviances are not significantly different from each other, which makes the model with the most degrees of freedom the best model that fits the data, which is Model 1 in Table 9.

CONCLUSION

My conclusion is to retain the null hypothesis, which is that for this data, school-based drug prevention programs are not effective, irrespective of the way the messages are delivered. I have made the case as strong as I could, using homogeneity analyses, where variables are scaled to construct a more reliable and more global scale, while also enhancing the validity of the measurements. And by using multilevel analyses, executed at the proper level, that of the student, and by using triangulation and descriptive statistics to underscore the findings.

Given the danger of ecological fallacy, it is not surprising that the result of Hansen and Graham's (1991) analyses with the same data set differ from the ones reported here. Based on their class-level analysis, they report "that a *p* value of 0.0011 indicates a significant reduction in onset (of alcohol use) attributable to normative education" (last cited p. 414). The fallacy in this conclusion may be twofold. The analyses were executed at class level and can only lead to the conclusion that classes receiving normative education have, on average, a reduction in onset of alcohol use, which may be attributable to NORM. The other fallacy is that causal statements are hard to defend when existing classes are used. I know that the method used by Hansen and Graham (1991) was the usual way to analyze data in drug prevention research (e.g., Dukes, Ullman, and Stein 1995). One of the reasons is that reviewers systematically rejected papers based on student-level analyses, out of concern for intraclass correlation, thereby ignoring that results of aggregated analyses are not necessarily the same as the ones obtained from student-level analyses. The two articles by Dukes, Ullman, and Stein (1995,

1996) are another example of differences of results due to a change in level of analysis. The student-level analysis (last cited 1996) does not show program effects, which contradicts their earlier findings (1995) which was based on class-level analyses.

The different results between the analysis in this article and Hansen and Graham's (1991) analysis can also be explained by the use of a different response variable. Hansen and Graham dichotomized the responses of three items, the lifetime alcohol use (Item 19) and alcohol use in the past month (Item 20) and in the past week (Item 21). The code 1 is used to indicate alcohol use (from half a drink to more than 11 drinks; see Appendix for categories of the variables), and 0 for no use, including a "sip." After this data reduction, the student data are aggregated to class level. It is not surprising that this choice of response variable, as compared to the composite score used in the analyses presented here, leads to different conclusions regarding program effects. Knowing what choice of analysis level and what choice of response variable will yield the most reliable results can only come from looking at the consequences of these choices. I have reported why I have chosen the scaling technique homogeneity analysis and the multilevel analysis method. I have also reported exploratory analyses and compared them with the results of more traditional techniques.

All my results seem to point in the direction of zero effect of NORM where neither statistically significant effects nor important effects are found. The exploration of a moderator effect of NORM was not successful, nor did NORM significantly lower the number of abstainers. Although the percentage of abstainers from 1987 to 1988, calculated over program conditions, shows that NORM has the lowest number of students changing from abstainers to alcohol users, and RT shows the usual high numbers, the difference is neither large enough to be statistically significant nor important. The same result is obtained in a comparison using Item 35 (see Appendix), where the question is asked how many times the respondent has been drunk. An increase is observed over the year in the number of students that answer yes to this question. And indeed, this number is largest in the control group and smallest in the NORM group, but again the difference is too small to reach statistical significance. The question if such a small number is an important difference is an economical one.

After finding no effects of drug prevention programs, the rich data set is further explored in search for student risk factors in relation to alcohol use. "At risk behavior" of students is defined here as "the action of a person or the environment that raises the risk for future alcohol abuse." This concept is used in the drug prevention literature, and programs are developed to counter this

risk or, at least, lower the risk. My analyses indicate that drinking in the environment of the respondent by friends and adults, who often drink to excess, is related to respondents' alcohol involvement and makes these students at risk. The interactions of NORM with friends drinking and respondents drinking are constructed to test if NORM has a moderator effect, by lowering these relationships. But again no moderator effects of NORM are found. If drug prevention programs fail, it may be because the influence of the environment is stronger than the influence of cognitive lectures and exercises. My analyses' results support the findings of Kandel (1974) that the example set by parents and peers is a crucial factor in drug use.

Other risk factors are present in these data. In analyses not reported here, I found that "trouble" in school, low grades, and "rebelliousness" (see Kreft 1996) are factors related to alcohol use. The literature mentions risk factors related to parental behavior, such as a bad or indifferent relationship between parents and friends of the child, or between low degree of parental guidance and/or low degree of parental trust. In this data set, after controlling for pretest use, no such effects are found to be significantly related to alcohol use (Kreft 1996).

Because causal factors of students' involvement with alcohol cannot be determined by data analyses, without strong support of a theory, the conclusions drawn from my analysis can only be used to exclude, not to include. My analyses support the conclusion that we can exclude that the two school-based drug prevention programs, RT and NORM, work.

NOTES

1. D.A.R.E. is the copyrighted acronym for Drug Abuse Resistance Education. It is administered by D.A.R.E., a nonprofit organization based in Los Angeles, California. The program is administered starting in the last grades of elementary school. Police officers are trained to teach students to resist drug offers, and instead accept a drug-free lifestyle. This program resembles the earlier mentioned RT program.

2. See Donaldson et al. (1995).

3. Resistance training is designed to help kids see the kinds of pressure to use drugs by teaching them skills to resist such pressure without losing friends. The program is based on the assumption that kids want to resist drug offers but simply lack the proper skills.

4. NORM is a normative program developed by Hansen and Graham (1991), based on social theory of Bandura (1977 and 1986), and Jessor & Jessor (1977). It is expanded from a single session in the project smart curriculum and based on the assumption that students overestimate prevalence and acceptability of alcohol and other drug use. By correcting overestimation, it is reported that the program is successful in lowering the onset of alcohol use in teenagers or in a reduction of use.

APPENDIX

HOMALS CATEGORY QUANTIFICATION TABLES

Item 19: How many drinks of alcohol have you had in your whole life?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
0.64	0.70	1 none	1,002	743
0.42	0.60	2 only a sip (for religious purposes)	355	304
0.18	0.36	3 only a sip (not for religious purposes)	760	756
-0.25	0.05	4 part or all of a drink	226	240
-0.52	-0.21	5 2 to 4	227	280
-0.97	-0.56	6 5 to 10	159	224
-1.59	-1.02	7 11 to 20	129	178
-2.23	-1.62	8 21 to 100	103	200
-3.07	-2.75	9 more than 100	50	105
-0.31	-0.74	missing	16	17

Item 20: How many drinks of alcohol have you had in the past month?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
0.35	0.44	1 none	2,383	2,209
0.09	0.32	2 only sip (for religious purposes)	130	100
-0.90	-0.58	3 only a sip (not for religious purposes)	232	243
-1.70	-1.06	4 part or all of a drink	99	159
-2.25	-1.67	5 2 to 4	115	161
-3.18	-2.18	6 5 to 10	35	93
-4.13	-2.83	7 11 to 20	13	36
-4.07	-3.63	8 more than 20	13	36
0.00	0.00	missing	7	10

Item 23: How many drinks of alcohol have you had in the past week?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
			0.23	
-0.19	0.05	2 only sip (for religious purposes)	67	63
-1.46	-1.20	3 only a sip (not for religious purposes)	116	96
-2.15	-1.72	4 1/2 or less	50	87
-2.44	-1.83	5 1	50	66
-3.35	-2.48	6 2 to 4	35	60
-3.55	-2.95	7 5 to 10	12	31
-4.24	-3.82	8 11 or more	11	22
0.00	0.00	9 missing	10	12

Item 24: How many days in the past month have you had alcohol to drink?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
0.32	0.41	1 none	2,586	2,402
-1.32	-0.97	2 1	240	286
-2.27	-1.72	3 2 to 3	117	198
-3.06	-2.17	4 4 to 7	48	90
-4.08	-3.09	5 8 to 14	14	23
-3.32	-3.43	6 15 to 30	12	29
0.00	0.00	7 missing	10	19

Item 25: How long has it been since you had any alcohol to drink?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
-1.88	-2.12	1 less than 24 hours	47	55
-2.25	-1.95	2 > a day, < a week	147	197
-1.39	-1.14	3 > a week, < a month	269	386
-0.36	-0.13	4 > a month, < 6 months	438	524
0.09	0.21	5 > 6 months, < a year	302	282
0.29	0.45	6 > a year	567	604
0.61	0.69	7 I never had any alcohol	1,243	977
0.00	0.00	8 missing	14	22

**Item 28: Think of the day during the past month when you drank the most alcohol.
How many drinks did you have that day?**

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
0.58	0.65	1 I never drink	1,459	1,256
0.05	0.22	2 no alcohol past month	766	863
-0.41	-0.29	3 sips	384	328
-1.10	-0.94	4 1	160	210
-1.81	-1.40	5 2	82	109
-2.02	-1.68	6 3	51	77
-2.15	-2.04	7 4	31	45
-2.78	-2.36	8 5 or more	81	140
0.00	0.00	9 missing	13	19

(continued)

APPENDIX Continued

Item 29: How often do you imagine yourself having a drink of alcohol?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
-1.88	-2.18	1 <i>often</i>	59	77
-1.61	-1.47	2 <i>sometimes</i>	232	337
-0.39	-0.18	3 <i>hardly ever</i>	819	889
0.42	0.48	4 <i>never</i>	1,905	1,726
5 <i>missing</i>	12	18		

Item 32: Do you think you will drink alcohol in the next couple of months?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
-2.79	-2.17	1 <i>yes</i>	117	223
-1.37	-1.05	2 <i>probably</i>	303	403
-0.38	-0.14	3 <i>I don't think so</i>	536	639
0.46	0.57	4 <i>no</i>	2,067	1,756
		<i>missing</i>	4	26

Item 33: Do you think you will ever drink alcohol every day?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
-1.55	-2.19	1 <i>yes</i>		28
-1.78	-1.25	2 <i>probably</i>	75	71
-0.93	-0.92	3 <i>I don't think so</i>	349	418
0.19	0.22	4 <i>no</i>	2,566	2,496
		<i>missing</i>	9	25

Item 34: Do you think you will ever drink alcohol every month?

<i>Pretest</i>	<i>Posttest</i>	<i>Original Code and Wording</i>	<i>N 1987</i>	<i>N 1988</i>
0.39	0.47	1 <i>no</i>	2,195	1,915
-0.55	-0.32	2 <i>I don't think so</i>	452	585
-1.23	-1.17	3 <i>probably</i>	293	381
-2.03	-1.98	4 <i>yes</i>	114	14
		<i>missing</i>	9	26

Item 35: How many times have you ever been drunk?

Pretest	Posttest	Original Code and Wording	N 1987	N 1988
0.28	0.35	1 <i>never</i>	2,520	2,340
-0.72	-0.47	2 <i>only once</i>	287	298
-1.94	-1.31	3 <i>2 to 4 times</i>	140	253
-2.42	-2.17	4 <i>5 to 10 times</i>	43	75
-3.43	-2.82	5 <i>11 to 20 times</i>	14	29
-4.34	-3.52	6 <i>> 20 times</i>	15	28
0.00	0.00	7 <i>missing</i>	8	24

Item 38: Do you think you will get drunk in the next couple of months?

Pretest	Posttest	Original Code and Wording	N 1987	N 1988
0.23	0.33	1 <i>no</i>	2,676	2,464
-1.29	-0.96	2 <i>I don't think so</i>	211	348
-2.42	-1.94	3 <i>probably</i>	85	136
-3.14	-3.11	4 <i>yes</i>		41
		<i>missing</i>	14	28

REFERENCES

- Aitkin, M. A. and N. Longford. 1986. Statistical modelling in school effectiveness studies. *Journal of the Royal Statistical Society* 149A:1-43.
- Bandura, A. 1977. *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- . 1986. *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Barcikowski, R. S. 1981. Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics* 6 (3): 267-85.
- Benzecri, J. P. 1973. *Analyse des donnees* (Data analysis). Paris: Dunod.
- Brown, Joel H. and Marianne D'Emidio Caston. 1995. On becoming "at risk" through drug education: How symbolic policies and their practices affect students. *Evaluation Review* 19:451-91.
- Burstein, Leigh. 1980. Assessing differences between group and individual level regression coefficients: Alternative approaches. *Sociological Methods & Research* 7:5-27.
- Carbonari, Joseph P., Philip W. Wirtz, Larry R. Muenz, and Robert L. Stout. 1994. Alternative analytical methods for detecting matching effects in treatment outcomes. *Journal of Studies on Alcohol* (Suppl. 12): 83-90.
- D'Emidio-Caston, Marianne, and Joel H. Brown. 1998. The other side of the story: Student narratives on the California Drug, Alcohol, and Tobacco Education Programs. *Evaluation Review* 22 (1): 93-115.
- De Leeuw, Jan and Ita G. G. Kreft. 1986. Random coefficient models for multilevel analysis. *Journal of Educational Statistics* 11 (1): 57-86.
- Donaldson, Stewart I., John W. Graham, and William B. Hansen. 1994. Testing the generalizability of intervening mechanism theories: Understanding the effects of adolescent drug use prevention interventions. *Journal of Behavioral Medicine* 17:2.

- Donaldson, Stewart, I, John W. Graham, Andrea M. Piccinin, and William B. Hansen. 1995. Resistance-skills training and onset of alcohol use: Evidence for beneficial and potential harmful effects in public schools and in private Catholic schools. *Health Psychology* 14:291-300.
- Dukes, Richard L., Jodie B. Ullman, and Judith A. Stein. 1995. An evaluation of D.A.R.E. using a solomon four-group design with latent variables. *Evaluation Review* 19 (4): 389-409.
- . 1996. Three-year follow-up of drug abuse resistance education (D.A.R.E.). *Evaluation Review* 20 (1): 49-66.
- Ennett, Susan T., Nancy S. Tobler, Christopher L. Ringwalt, and Robert L. Flewelling. 1994. How effective is drug abuse resistance education? A meta analysis of the project D.A.R.E. outcome evaluations. *American Journal of Public Health* 84 (9): 1394-401.
- Graham, John W., Linda M. Collins, Stuart E. Wugalter, N. K. Jacob Chung, and B. Hansen. 1991. Modeling transitions in latent stage-sequential processes: A substance use prevention example. *Journal of Consulting and Clinical Psychology* 59 (1): 48-57.
- Graham, John W., G. S. Marks, and William B. Hansen. 1991. Social influence processes affecting adolescent substance use. *Journal of Applied Psychology* 76:291-8.
- Greenacre, M. J. 1984. *Theory and applications of correspondence analysis*. London: Academic Press.
- Guttman, Louis A. 1941. The quantification of a class of attributes: A theory and a method of scale construction. In *The prediction of personal adjustment*, edited by P. Horst. New York: SSRC.
- Hannan, M. T. 1974. Problems of aggregation. In *Causal models in the social sciences*, edited by H. M. Blalock, chap. 27, 473-508. Chicago: Aldine.
- Hansen, W. B. 1993. School-based alcohol prevention programs. *Alcohol Health & Research World* 17 (1): 54-60.
- Hansen, W. B. and J. W. Graham. 1991. Preventing alcohol, marijuana, and cigarette use among adolescents: One-year results of the Adolescent Alcohol Prevention Trial. *Preventive Medicine* 20:414-30.
- Hensen, W. B., John W. Graham, Bonnie H. Bolkenstein, and Louise Ann Rohrbach. 1991. Program integrity as a moderator of prevention program effectiveness: Results for fifth-grade students in the adolescent alcohol prevention trial. *Journal of Studies on Alcohol* 52 (6): 568-79.
- Hedeker, D., R. D. Gibbons, and J. M. Davis. 1991. Random regression models for multicenter clinical trials data. *Psychopharmacology Bulletin* 27:73-7.
- Hedeker, Donald, Robert D. Gibbons, and Brian R. Flay. Forthcoming. Random-effects regression models for clustered data: With an example from smoking prevention research. *Journal of Consulting and Clinical Psychiatry*.
- HOMALS. 1989. *SPSS-categories*. Chicago: Author.
- Jessor, R. and S. L. Jessor. 1977. *Problem behavior and psychosocial development*. New York: Academic Press.
- Kandel, D. B. 1974. Interpersonal influences on adolescent illegal drug use. In *Drug use: Epidemiological and sociological approaches*, edited by E. Josephson and E. E. Carroll, 207-39. Washington, DC: Hemisphere.
- Kreft, Ita G. G. 1994. Multilevel models for hierarchically nested data: Potential applications in substance abuse prevention research. In *Advances in data analysis for prevention intervention research*. NIDA Monograph #142, edited by Linda M. Collins and Larry A. Seitz, 140-84. Rockville, MD: National Institutes of Health.

- . 1996. *The evaluation of the effects of drug prevention programs on junior high school students' alcohol consumption with homogeneity and multilevel analyses*. PERC Report #1, Charter School of Education, California State University, Los Angeles.
- . 1997. The interactive effect of alcohol prevention programs in high school classes. In *New methodological developments in prevention: Alcohol and substance abuse research*, edited by Kendall, J. Bryant, Michael Windle, and Stephen West. Washington, DC: American Psychological Association.
- Kreft, Ita G. G. and Edith D. De Leeuw. 1988. The seesaw effect: A multilevel problem? *Quality and Quantity* 22:127-37.
- Kreft, Ita G. G., Jan de Leeuw, and Rien van der Leeden. 1994. Review of five multilevel analysis programs: BMDP-5V, GENMOD, HLM, ML3, VARCL. *American Statistician* 48 (4): 324-35.
- Laird, N. W. and J. H. Ware. 1982. Random-effects models for longitudinal data. *Biometrics* 38:963-74.
- Longford, N. T. 1987. A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika* 74:817-27.
- May, Carl. 1993. Resistance to peer group pressure: An inadequate basis for alcohol education. *Health Education Research: Theory and Practice* 8 (2): 159-65.
- ML3. 1989. Software for two- or three-level analysis written by Rabash, Institute of Education, University of London.
- Raudenbush, S. and A. S. Bryk. 1986. A hierarchical model for studying school effects. *Sociology of Education* 59:1-17.
- Robinson, W. S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review* 15:351-7.
- SPSS. 1989. *SPSS-Categories*. Chicago: SPSS, Inc.
- Van de Geer, John P. 1993a. *Multivariate analysis of categorical data: Applications*. Advanced Quantitative Techniques in the Social Sciences Series, #1. Newbury Park, CA: Sage.
- Van de Geer, John P. 1993b. *Multivariate analysis of categorical data: Theory*. Advanced Quantitative Techniques in the Social Sciences Series, #2. Newbury Park, CA: Sage.
- Wang, C. 1993. *Sense and nonsense of statistical inference*. New York: Marcel Dekker.

Ita Kreft works at California State University, Los Angeles in the School of Education. She teaches statistics and research methodology. Her research interest is in the application of multilevel analysis to social research problems; this includes analysis of data where individuals are studied in context, with special emphasis on the influence of the individual, as well as the influence of context on human behavior. Her special interest is in drug prevention research and the social and individual characteristics that leads to substance abuse as exemplified by her publications and presentations at National Institute on Drug Abuse (NIDA).

The evaluation of community-based programs poses special design and analysis problems. The present article focuses on two major types of errors that can occur in such evaluations: false positives—incorrectly declaring a program to be effective—and false negatives—incorrectly declaring a program to be ineffective. The evaluation of a national demonstration of community-based programs to reduce substance abuse, Fighting Back, is used to illustrate several approaches to reduce the probability of errors. Both those errors that are affected by the design and those by analytic approaches are considered. Ways to assess multiple outcomes and to match the complexity of the program with design and analytic strategies are proposed. Community trials are complex interventions, and, although they can provide very useful information, their outcomes have to be understood in terms of the constructs they test and the contexts within which they are carried out.

ZERO EFFECTS IN SUBSTANCE ABUSE PROGRAMS

Avoiding False Positives and False Negatives in the Evaluation of Community-Based Programs

DAVID RINDSKOPF

LEONARD SAXE

City University of New York

As noted in the introduction to this special issue, a depressing discordance exists between claims made about programs to prevent and treat substance abuse and methodologically sound data to support these contentions (see Kreft 1998 [this issue]). The disparity between the optimism of program developers and the results found by some evaluators is, perhaps, understandable and represents the hope that thoughtful interventions can redress the serious problems created by substance abuse. Remediation programs, particularly community-based substance abuse initiatives that have broad goals (see Aguirre-Molina and Gorman 1996; Winick and Larson 1997), engender a host of methodological challenges. The complexity of community trials makes it difficult to resolve questions about program effects (cf. Connell, Aber, and Walker 1995; Murray, Moskowitz, and Dent 1996), and

it is not surprising that optimistic assessments of the potential of these efforts has not been matched by supporting data (Feinleib 1996; Susser 1995). The present article describes several key design and analysis issues inherent in the evaluation of substance abuse programs and suggests some practical solutions to these problems. The goal is to suggest ways in which systematic evaluation can be used to promote development of effective substance abuse programs.

The *raison d'être* for program evaluation is decision making, and the objective of an evaluative research study is to provide information that can improve social planning. A program can be truly effective or not, and it can be declared effective or not. The combination of two true and two declared states of effectiveness results in a 2×2 table, similar to that used in the familiar hypothesis-testing framework, containing four possible outcomes. Our goal is to avoid the two outcomes that result in errors: (a) deciding that a truly ineffective program is effective (false positive, corresponding to a Type I error) and (b) deciding that a truly effective program is ineffective (false negative, corresponding to a Type II error). The focus of this article is how to design program evaluation studies and conduct statistical analyses of their findings that minimize the probability of each type of error.

The discussion of how to deal with potential errors is organized around a description of the design and analysis of a large national substance abuse demonstration program, Fighting Back. Fighting Back was developed by the Robert Wood Johnson Foundation (RWJF) to test the proposition that demand for illicit drugs and alcohol can be reduced by organizing communities to change attitudes and norms about substance use (Jellinek and Hearn 1991). The foundation was committed not only to providing support for communities to develop demand reduction efforts but also to developing knowledge that could shape substance abuse policy. Knowledge generation was a primary justification for their investment of nearly \$100 million during more than 10 years. Foundation officials are not disinterested observers and would be disappointed if program effects were not found, but finding the truth is more important than rationalizing a failure.

Community-based interventions such as Fighting Back present a host of methodological challenges for evaluation (Weiss 1995; Winick and Larson 1997). One inherent problem in such programs is that there may be substantial variation in how the construct is implemented across communities. The variation makes it more difficult to detect program effects and leads to potential false negatives. At the same time, the broad focus of these programs often leads to a large number of outcomes that the programs would like to

affect. The number of outcomes tested increases the probability of significant effects by chance and, thus, can produce false positive conclusions.

General issues that affect the conduct and interpretation of evaluative studies of broad-based community intervention programs are discussed below. The discussion begins with a summary of the Fighting Back evaluation (see also Saxe et al. 1997). Issues directly related to the evaluation process are then presented, and the discussion is broadened to include issues not encountered in Fighting Back but that could plausibly be expected to occur in the evaluation of such programs. Although in some instances we can merely identify problems, in other cases—most notably, the problem of multiple statistical tests—solutions are suggested.

FIGHTING BACK

The objective of Fighting Back was to demonstrate the feasibility of reducing substance abuse through comprehensive and coordinated community efforts (Jellinek and Hearn 1991; Spickard, Dixon, and Sarver 1994). The program was developed on the premise that reducing substance abuse, and consequent harm, requires development of a system of prevention, treatment, and aftercare. A basic assumption of Fighting Back is that substance use and abuse are influenced by physical and social environments (Kadushin et al. forthcoming). In contrast to the traditional focus of alcohol and other drug (AOD) programs on supply reduction (Gorman 1993; Humphreys and Rappaport 1993), Fighting Back has focused on changing the environments that promote and sustain the demand for AOD. Because the environment affects AOD use in multiple ways, the implicit theory is that entire communities must organize and collaborate to address the problem.

The essential element of each Fighting Back initiative has been the development of a community-wide common vision that would foster collaborative efforts to address the substance abuse problem. The implicit theory (see Jellinek and Hearn 1991) is that broad-based partnerships, involving a community's major constituents, can bring multiple perspectives to bear on substance abuse problems, enhance feelings of ownership of AOD problems, and increase motivation to promote and sustain both treatment and prevention efforts. The architects of Fighting Back were aware that long-standing community traditions and interorganizational tensions might inhibit development of community-wide systems and noted, "The principal barriers are essentially political" (Jellinek and Hearn 1991, 78).

Fighting Back is RWJF's flagship alcohol and drug prevention program and has been implemented in 14 communities (typically a small city or portion of a city). Although the foundation believed that effective solutions to AOD problems share common elements, they recognized the uniqueness of each community. The communities differ demographically in how they experience AOD problems; in the resources, institutions, and procedures available to address substance abuse; and, perhaps most important, in their history of collaborative working relationships. Thus, individual programs were given considerable latitude to implement the global concepts.

Developing a "shared vision" and a "coordinated effort to change" was operationalized by the requirement that community-wide systems be both broad and deep, spanning the community both vertically (from community elites to grassroots activists) and horizontally (reaching across the various political, business, and service domains). RWJF mandated sites to create a Citizen's Task Force representing all the key constituents of a community—business, clerical, medical, legal, and neighborhood—to oversee the program. Once implemented, day-to-day activities were to be carried out by Fighting Back staff, under the direction of an executive committee, which was also to represent the multiple interests of the community.

The Fighting Back demonstration assumed that there are multiple causes of substance use and abuse and that effective programs need to address a comprehensive set of domains. It was assumed that the program's impact would be broad-based and would decrease the use and abuse of alcohol and illicit drugs, along with the harms caused by substance abuse. Thus, a key evaluation issue has been to record what programs communities were able to mount and to understand how communities achieved synergy among programs. The ultimate issue, however, is whether the program affected the actual rate of substance use and the harms associated with abuse.

Fighting Back was designed as a 7-year demonstration, with the first 2 years devoted to planning (c. 1991-1992). It has been implemented in 14 communities, 12 of which have been the focus of the present evaluation (see Saxe et al. 1997). Each Fighting Back site received about \$3 million during a 5-year period. Prior to the end of the 5-year implementation phase (mid-1997), the foundation decided to extend the program up to 5 additional years. Only the eight sites considered to have made the most progress are eligible for continued funding. Nevertheless, all 12 will continue to be part of the program (including eligibility for technical assistance) and are part of the evaluation.

EVALUATION DESIGN

The evaluation of the Fighting Back program is designed as a quasi-experiment, with each treatment community compared to multiple matched comparison sites on several outcome dimensions of use, harm, and attitudes. The nested structure of the design calls for a multilevel analytic framework and is being used to test several types of outcomes (see Rindskopf et al. 1997). These outcome measures include (a) a multiwave survey to assess AOD use, as well as attitudes and perceptions of AOD use, crime, and community issues, and (b) indicators over time of crime, traffic accidents, morbidity, and mortality. These data are collected both from Fighting Back and comparison sites. For the survey, our design has people nested within communities and communities nested within matched sets (usually a state). For the indicators, yearly observations are nested within sites and sites are nested within matched sets.

Because evidence of change in AOD use and attitudes will not necessarily validate (or invalidate) the program concept (see Hunt 1994; Lorion 1994), assessment of the Fighting Back construct is dependent on the extent to which changes in AOD use and attitudes can be linked to program features. Thus, the design includes an assessment of the community structures and program strategies used to combat substance use and abuse and the subsequent harm that results. Evaluation data collected during the past 3 years (since the present evaluators assumed responsibility) document how participating communities fostered development of a broad array of strategies to deal with substance abuse. These data provide support for the feasibility of the program construct and, along with outcome data, establish a baseline for interpreting future results (Saxe et al. 1997). These data are supplemented by intensive ethnographic studies of a sample of Fighting Back sites. The ethnographic studies describe how and to what extent the concepts of Fighting Back were put into practice.

Because the evaluation is not nearly complete (nor will it be for several years), it is not possible to say whether Fighting Back worked or did not—even if that were a possible question to answer (see Granger 1997). Preliminary evidence from at least one aspect of the study, however, disappointed some proponents of the program. In the first wave of survey data, as well as the analysis of indicators, no outcome variable shows even a moderately large program effect (see Kadushin et al. forthcoming; Saxe et al. 1997). Although the finding was, in some respects, reassuring because it validated the assumption of a baseline, for many communities it was discouraging feedback on their efforts to address substance abuse problems.

The discussion below focuses on several research design issues that we tried to anticipate in order to avoid misinterpretations of the findings. Several issues that might lead to a conclusion of zero effects (either true or due to error) or nonzero effects (again, either true or in error) are discussed. Although solutions to some problems have been identified, with others it is only possible to point out the problem's existence.

REDUCING THE RISK OF FALSE NEGATIVES AND POSITIVES

Reducing the risk of error, in general, requires a conceptual understanding of what is being attempted. It requires, as well, realistic expectations about outcomes.

Global (general) versus local effects and evaluation. Community-based programs, even those that share a similar focus and strategy, are often implemented in very different ways across communities (cf. Connell, Aber, and Walker 1995). Although Fighting Back staff made it clear what the expected outcomes were, in general terms, each site had some of its own goals and objectives (e.g., one site focused on youth, another on specific neighborhoods). The national evaluation was intended to be a global, or general, evaluation, in which all sites were compared on the same outcome variables. The Fighting Back theory was that each community would choose approaches and methods unique to its own problems and resources, but that all sites would have the same objectives. But such guidelines give communities considerable latitude in how they address AOD problems. One problem is that some communities might focus on a narrow set of AOD problems in the community (e.g., care of addicted newborns), while not addressing the main Fighting Back issues or resulting in outcomes measured by the global evaluation we are conducting. The diversity in target outcomes makes the evaluation difficult and replication of the program nearly impossible. Further, it takes detailed field work to determine whether anything happened at each site, and, if so, what activities are directly related to the Fighting Back intervention. A nearly impossible task would be to determine how to repeat in other communities any successes that were found.

The danger of missing real effects because they are not measured by our global indicators is being mitigated, at least, by encouraging communities to collect their own local evaluation data. Each site can thus target measures that are appropriate for its local programs, over and above the measures for the

global evaluation. From the global evaluation standpoint, these will supplement, and not supplant, the broader analyses.

The global versus local problem manifests itself in more ways than whether outcome measures that are common across sites can be used. An additional problem concerns the target population. Site boundaries do not necessarily follow political boundaries, and the target area of each site had to be painstakingly mapped and related to zip codes, police precincts, school districts, and other political units. All of our data collection, in particular surveys and indicators, were designed to collect data only from the target areas (see Beveridge et al. 1997). From the national/global evaluation standpoint, the entire population within those boundaries is the target population. If some aspects of the program are targeted at a subpopulation or subsection within that area, then the effect will be diluted. This might cause the evaluation to miss an effect that is real because the effect was not large enough with respect to the national/global evaluation's target population. From the standpoint of the global evaluation, this is as it should be; locally, however, each site has to conduct its own evaluation to demonstrate that any such effects exist, or risk what would be (to the local sites) a false negative finding.

One statistical approach that might be taken when each site has its own focus is to conduct a meta-analysis (cf. Cooper and Hedges 1994), with sites rather than studies being the unit of analysis. If effects from all sites are placed on a common scale, as is typically done in meta-analysis, then an overall analysis could examine the "average" treatment effect, as well as variations in effect across sites. Furthermore, if there is variation across sites in outcome, one can include site-level characteristics to see which are associated with larger effects (see discussion below). Such a strategy would be particularly important in large-scale programs with varied outcomes where only a small common core of outcomes (or none) would be appropriate. The meta-analytic strategy reduces the likelihood of missing real effects that could occur if the common core of outcomes is so small that most important outcomes would go unmeasured.

Great expectations I: Program size. It is not surprising, given the foundation's investment in the program (the largest of any of their demonstration efforts), that they have high expectations. To justify a long-term investment in these communities, a credible case had to be made that the expenditure was worthwhile. Yet, as costly as the program is, each site receives less than \$1 million per year, and their foundation funding represents only a small proportion of their annual expenditures on AOD programs. Although the theory is that this "small" amount can be leveraged to produce large effects by increasing the organizational capacity of the site, there are many reasons why such leveraging might not happen or, perhaps, might be much smaller

than hoped for. An actual zero effect is a risk for the program itself, not for the evaluation. If the program theory is wrong, and the effects are truly zero, then that is what will be found. But nonzero effects that are much smaller than expected increase the risk of false negative findings, as discussed in more detail below.

Great expectations II: Effect size and power. How big an effect should a program have? On one hand, program advocates expect that their program will have a large effect. Funders must also believe it, or they would not provide the funds. On the other hand, a complex problem might be expected to be resistant to massive change. It is interesting to note that one of our most successful health care prevention efforts, the effort to reduce tobacco consumption, has resulted in a large long-term change that is barely visible if one compares rates year by year. Although the number of current adult smokers has been reduced since 1965 by nearly 20%, in no single year does the rate change exceed 1% (Centers for Disease Control 1997).

If one believes that the effect size will be small during the time of the study, then enormous sample sizes might be needed to detect it. As an example, if one believes that you can only get 5% of drug users to quit, and if the rate of drug use is 10% in your target population, then you want to reduce the usage from a rate of .100 to a rate of .095. The sample size to have a power of .80 to detect that amount of change, with $\alpha = .05$, is more than 55,000 per group. The results of such a power analysis are enough to turn people into strong defenders of the view that their program will have a large effect. Nevertheless, when it comes to an evaluation, they may be doomed to miss a real (though small) effect because of low statistical power.

Great expectations III: Time lag. Given the large investment in the program, it is also not surprising that there were high expectations for relatively quick effects. Most evaluators believe that their task is a lot harder if they are not brought into a project early; however, it may be even more important to maintain the evaluation after the program has ended, if that is how long it takes for the effects to occur. As an obvious example, one would not expect alcoholic dementia to decrease appreciably for many years, even if a program to combat alcohol abuse were very successful. But leaving aside these obvious cases, one should expect sluggishness in the transmission of effects in a complicated system. It may take a year or more to organize all of the agencies interested in AOD to start planning how to coordinate their efforts, and several years more before they can actually coordinate efforts. Bad timing in an evaluation can, in these cases, miss real effects that do not develop until after the evaluation.

In the present case, the qualitative (community studies) aspect of our evaluation is examining a sample of Fighting Back sites to determine how

slowly or quickly service coordination developed (see Jones and Fisher 1997). Obviously, if nothing has happened at a site, then there should be no effect to measure. From a program standpoint, this would be a poor outcome, but from the evaluation standpoint, it would provide a good baseline measure to assess the analytic strategy and our selection of control sites and variables.

One must guard, however, against the possibility of misuse of intermediate evaluation results. The evaluators could be treated as hostile by the remaining program sites if the intermediate results were used to reduce funding for sites that are not performing well. This puts the evaluators in a difficult position because they have responsibilities to the funding entity (to report useful information in a timely manner), as well as to the program sites (not to misuse the access to data that might not always show the site in a positive light). Although such problems can be anticipated, they cannot always be avoided, and evaluators should be wary of making promises that cannot be kept. "Premature evaluation" presents dangers to the validity of the evaluation. First, programs in what might become truly successful sites are at risk of being dropped before they have a chance to develop fully, thus missing true nonzero effects. Second, the remaining sites would be unrepresentative, having resulted from what some have called "creaming," and the program effects could then be overstated. Thus, both false negatives and false positives are possible consequences of premature evaluation.

Diffusion of treatment. The idea behind Fighting Back seemed so compelling to some that, even before it was fully implemented (and before it was evaluated), it was adopted by other organizations (see Aguirre-Molina and Gorman 1996; Winick and Larson 1997). Various forms of this idea have been implemented in a number of communities, many funded by the federal government under the Community Partnership Program (Kaftarian and Hansen 1994). The number of potential control sites was limited for our survey (because of the need to approximately match on size and demographics and stay within state boundaries where possible), so possible diffusion of the treatment into control sites was a serious concern. The situation would have been even more serious had the control sites been selected before these other programs started. Fortunately, many control sites still have no comparable programs. Our design also called for multiple control sites to be matched with each Fighting Back site; part of the reason for this was to ensure that we would still have a good chance of having true control communities by the end of the study. Some programs will not be so lucky: they will become so popular, even before an evaluation is completed, that they will be widely adopted. Evaluation data are obviously not the only force driving the acceptance of program ideas. In such situations, other quasi-experimental design strategies, probably involving time series or retrospective pretests, must be used. Obviously, the

diffusion of treatment increases the probability of false negatives because successful programs in control sites will reduce or eliminate the treatment-control differences.

Multiplicity of outcomes and the need for multiple hypothesis tests. Programs such as Fighting Back can legitimately claim that they are broad, that they have multiple goals, and thus should be measured on multiple outcomes. Assessing multiple outcome is, however, methodologically complex, because the more statistical tests that are done, the larger the probability of at least one being significant even if there are uniformly zero effects across all sites and all outcome measures. Tukey (1977) discussed this problem in the context of clinical trials, which is closely related to evaluation of community substance abuse initiatives.

There are a host of ways to deal with the multiple outcome issues. The ideas raised by Campbell (1966, 1978) on pattern identification are particularly useful. Campbell (1978) emphasized "the epistemic priority of patterns rather than particles" (p. 191) and that "qualitative, common-sense knowing of wholes and patterns provides the enveloping context necessary for the interpretation of particular quantitative data" (p. 192). Pattern matching can, in fact, strengthen one's confidence in the data.

Consider, for example, the obvious problem that out of 100 independent significance tests, 5 are expected to be significant even if the null hypothesis is true. But what if 20 out of 100 tests are significant? This is certainly much greater than would be expected if the null hypothesis were true. Furthermore, the direction of results must be considered: If all signs are in the direction of success (although not necessarily statistically significant individually), this is strong evidence of program effectiveness because if it were not, about the same number of negative as positive results would be expected. This suggests, at a minimum, the use of simple nonparametric tests (e.g., sign tests based on the binomial distribution) to test the pattern of results.

One can also use parametric versions of this procedure. For example, suppose one calculates a large number of standardized effect sizes. A half-normal plot of these would be a straight line if there were no true nonzero effects. If the largest effects do not lie on a straight line but, instead, are well beyond expectation, one could conclude that the effects are real in spite of being few in number. Fienberg (1980) describes an example of this technique in the context of sorting through a large number of estimated effects in loglinear models.

In some circumstances, the use of multivariate rather than univariate tests can reduce the total number of statistical tests that are done. A multivariate analysis of variance (MANOVA) conducted on 10 dependent variables may be more useful than 10 separate analyses of variance (ANOVAs), but only if

there are relatively strong relationships among those 10 variables; MANOVA will only reduce the dimensionality of the problem under those circumstances. Furthermore, in evaluations such as *Fighting Back* that involve categorical outcomes (e.g., drug dependency status), no good analogue to MANOVA is available.

Another approach to the multiplicity problem is to divide outcomes into two classes: those that should be affected by the program and those that should not. If the "relevant" outcomes are the ones that show significance, whereas the "nonrelevant" outcomes do not, multiple tests of significance are not problematic. The basic idea for this comes from the "Nonequivalent Dependent Variables Design" of Cook and Campbell (1979, 97, 118 ff; see also Rosenbaum 1995). Cook and Campbell describe a hypothetical study in which a school implemented a change in their mathematics curriculum stressing algebra over geometry and arithmetic. If gains are made in algebra but not in geometry or arithmetic (or to not as great a degree), then the evidence of a program effect is much greater than if there were no other related outcome variables measured.

Some examples of this use of what might be called "control" or "placebo" dependent variables exist in the evaluation and policy literature. Thus, for example, to investigate the effect of the introduction of television on reading, Parker et al. (1971, cited in Cook and Campbell 1979) examined the number of library books checked out each year. They found that, when television was introduced, the number of fiction works checked out declined greatly but not the nonfiction works. Thus, nonfiction works served as a control outcome variable.

A more involved version of this approach was successfully applied in the Kansas City Preventive Patrol (KCPP) evaluation (Kelling et al. 1974; see also Saxe and Fine 1981). Kansas City was divided into 15 police patrol beats; 5 received normal levels of police patrol, 5 received a much higher than normal level of patrolling officers, and 5 received a much lower than normal level of patrol. The change in level of each of a large number of crimes was followed. With a large number of significance tests, one might expect at least some to be significant. But the evaluators could separate the outcomes into crimes that should be affected by increased patrol (e.g., burglary) and those that should not (e.g., homicide). The results showed a clear pattern in that most of the crimes that should be affected had the expected outcome that higher levels of patrol were associated with greater drop in crime, whereas the crimes for which there was no such expectation showed no discernible relationship of patrol level to crime reduction.

In the case of *Fighting Back*, the situation is not as straightforward as for the KCPP evaluation. It may be difficult to determine which measures are

theoretically influencible (although some are obviously not and become "placebo outcomes"). Furthermore, the outcomes may vary from site to site. As discussed above, one approach to this problem is to conduct a meta-analysis. To deal with multiple outcomes at each site, one could either have outcomes nested within sites or lump all outcomes from a site together by averaging.

Because outcome variables may have been only vaguely specified in advance of program implementation and evaluation, they may be subject to *post hoc* rationalization. Thus, for example, although the program may be clearly targeted to reduce substance abuse, there are a variety of ways to operationalize substance use and subgroups in which it can be measured. Timing is critical: If program staff can commit to specific goals before the program is implemented, the *post hoc* nature will be minimized. After program implementation, cognitive dissonance reduction is too easy; an artificial example would be, "We didn't really think we could reduce drunk driving accidents very much in such a short period of time." If, however, these can be specified in advance, then (as with preplanned tests in ANOVA) one need not worry as much. (Also, this would solve Tukey's [1977] problem with Bayesian approaches. One would have different prior distributions for different outcome variables, depending on whether the program should theoretically affect that outcome variable.)

The tactic of having some outcome variables that should *not* be affected by treatment is unusual; in fact, it is contrary to the usual evaluation goal, which is to select outcome variables that are sensitive to treatment effects. But often, for little extra cost, data are also available on at least some variables that should not be affected. In the present case, a number of existing indicators are being used. The additional cost to analyze a few variables that should not be affected by the Fighting Back program is a powerful control strategy. When random assignment is not possible, such placebo outcomes can be very informative (see Rosenbaum 1995).

Another important way to reduce chance effects due to the multiplicity of statistical tests is the use of multilevel statistical models (Bryk and Raudenbush 1992; Goldstein 1995). Without multilevel models, it is tempting to examine each site separately and, thus, conduct a large number of statistical tests. With multilevel models, one can reduce the number of statistical tests by having two global hypothesis tests for each dependent variable: (a) Is the overall (average) effect different from zero? (b) Is there variability across sites in the effect (i.e., is the site-to-site variation different from zero)? In the case of Fighting Back, seven major variables from the survey were examined (alcohol, marijuana, cocaine, and other illicit drug use; binge drinking; alcohol; and drug dependency). Had each of the 12 sites been examined

separately, the analysis would have involved 84 (12×7) significance tests just to examine the main effect of the program. Using multilevel models reduced the analysis to 7 tests of the average effect and 7 tests for variability in the effect across sites, for a total of 14 tests of program effect. The distribution of effect sizes and significance levels was what one would expect if there were no program effects.

Program implementation. A typical concern in evaluation is whether a program was implemented at all and, if so, how well. In medical research (which includes much research on drug programs), a similar distinction is made between the efficacy versus effectiveness. A treatment may seem to work (is efficacious) under the best test conditions (e.g., at a university hospital) but does not work (is ineffective) when implemented in ordinary practice. The situation for Fighting Back is complicated because, technically, Fighting Back is not a program at all. That is, Fighting Back has no set procedures, materials, or methods; in this respect, Fighting Back is more like a policy, though administered by a foundation rather than the government (for an analogous attempt by the federal government, see Yin and Kaftarian 1997, for a description of the Center for Substance Abuse Prevention [CSAP] evaluation). The program has a clear goal—to increase coordination among key members of the AOD community—but it can be accomplished by many means, and some will be more relevant in particular communities than others. Because it is difficult to know what Fighting Back should look like, and because it can take on a different form in each site, judging implementation is not a straightforward task. One major component of our evaluation, community studies, has been used to assess implementation.

If implementation is low or absent in one or two sites, then the program could still get a fair evaluation among the remaining sites. If, however, implementation is poor or absent at many or most sites, the only conclusion would be that the program is difficult (perhaps impossible) to implement with the level of support (both monetary and in technical assistance) provided. This would leave the granting agency with the unhappy task of deciding whether to increase funding (develop a more powerful program) or to acknowledge that the program, as designed, is not able to achieve the desired goals.

A relevant concept in the present case, and for many programs, is that of "leverage." Leverage is the idea that a small amount of resources, if applied correctly, can have a large effect. One can trace this idea back to the Greeks, but in modern times it goes back at least to Wiener's (1948) cybernetics, resurrected as chaos theory in the past decade, and to the related concepts from catastrophe theory (Zeeman 1977). Although such effects undoubtedly occur and could be taken advantage of, common sense suggests that such effects are rare. Many individual and societal problems are complex; an

empirical approach would point to the fact that these problems have existed for thousands of years and, if they were that easy to solve, they would have been solved already.

In terms of our focus on "false negatives and false positives," the implication is that if statistical analyses showed an effect but no program implementation was evident, one should not believe that there was an effect. No research design has the capacity to eliminate all possible threats to the validity of inferences about a program (cf. Campbell and Stanley 1966; Saxe and Fine 1981). To the extent that one can examine whether analytic results are consistent with the implementation, it is possible to be more certain about our conclusions. Conversely, if the analytic results conflict with the implementation evidence, one should be more cautious, to avoid declaring a "nonprogram" to be effective.

FINAL THOUGHTS

The evaluation of complex programs, such as efforts to address our society's substance abuse problems, offers many opportunities to make both major types of decision errors. It is possible to claim effectiveness for an impotent program by failing to control for plausible alternative hypotheses and by performing large numbers of statistical tests. Perhaps more important, it is possible to mistakenly declare effective programs ineffective, simply because one failed to measure the right variables or the right number or type of individuals or because of limited power. By examining patterns of results—both within and across studies—the likelihood of the first type of error can be reduced. In this article, a number of approaches have been offered to address this problem.

The Fighting Back evaluation has minimized many of the sources of the second type of error by ensuring that all Fighting Back sites agree on their target area, their target population (all residents of that area), and their desired outcomes. The sample size is adequate to ensure sufficient power to detect moderate size effects. If, at the end of the evaluation, one is left with zero effects, it is hoped that there will be little room to rationalize.

But one should not be sanguine that the results of the evaluation of Fighting Back, or any complex social experiment, will yield unequivocal results. In fact, the most common decision error may be to assume that a simple answer is possible to problems such as how a community should respond to substance abuse. As Weiss (1980) noted long ago, our concern should be with the accumulation of knowledge and we should allow for

"decision accretion." There is much that our evaluation studies can tell us about the character of communities and how they deal with substance abuse that can aid future efforts. It is critical that evaluators be able to tell the story unfettered by concerns about a need for unequivocal findings and that we employ design and analysis models that match the complexity of the programs being studied.

REFERENCES

- Aguirre-Molina, M. and D. M. Gorman. 1996. Community-based approaches for the prevention of alcohol, tobacco, and other drug use. *Annual Review of Public Health* 17:337-58.
- Beveridge, A., E. Tighe, M. J. Larson, D. Rindskopf, D. Livert, S. Weber, C. Swartz, J. McKenna, C. Ng, and L. Saxe. 1997. Monitoring archival indicators of alcohol and other drug harm: A Fighting Back progress report. Unpublished manuscript.
- Bryk, A. S. and S. W. Raudenbush. 1992. *Hierarchical linear models: Application and data analysis methods*. Newbury Park, CA: Sage.
- Campbell, D. T. 1966. Pattern matching as an essential in distal knowing. In *The psychology of Egon Brunswik*, edited by K. R. Hammond, 81-106. New York: Holt, Rinehart and Winston.
- . 1978. Qualitative knowing in action research. In *The social contexts of method*, edited by M. Brenner, P. Marsh, and M. Brenner, 184-209. New York: St. Martin's.
- Campbell, D. T. and J. C. Stanley. 1966. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Centers for Disease Control. 1997. CDC's tobacco information and prevention sourcepage. *CDC's Tobacco Info-Adult Prevalence*. [Online]. Available: <http://www.cdc.gov/nccdphp/osh/adstat1.htm>.
- Connell, J. P., J. L. Aber, and G. Walker. 1995. How do urban communities affect youth? Using social science research to inform the design and evaluation of comprehensive community initiatives. In *New approaches to evaluating community initiatives: Concepts, methods, and contexts*, edited by J. P. Connell, A. C. Kubisch, L. B. Schorr, and C. H. Weiss, 93-125. Washington, DC: The Aspen Institute.
- Cook, T. D., and D. T. Campbell. 1979. *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Cooper, H. M. and L. V. Hedges. 1994. *The handbook of research synthesis*. New York: Russell Sage.
- Feinleib, M. (1996). Editorial: New directions for community intervention studies. *American Journal of Public Health* 86:1696-8.
- Fienberg, S. E. 1980. *The analysis of cross-classified categorical data*. 2d ed. Cambridge, MA: MIT Press.
- Goldstein, H. 1995. *Multilevel statistical models*. 2d ed. New York: Halsted.
- Gorman, D. M. 1993. "War on drugs" continues in United States under new leadership. *British Medical Journal* 307:369-71.
- Granger, R. C. 1997. Establishing causality in evaluations of comprehensive community initiatives. Unpublished manuscript, Manpower Demonstration Research Corporation.

- Humphreys, K. and J. Rappaport. 1993. From the community mental health movement to the war on drugs: A study in the definition of social problems. *American Psychologist* 48:892-901.
- Hunt, G. P. 1994. Ethnography and the pursuit of culture: The use of ethnography in evaluating the Community Partnership Program. *Journal of Community Psychology*, CSAP special issue, 52-60.
- Jellinek, P. S., and R. P. Hearn. 1991. Fighting drug abuse at the local level: Can communities consolidate their resources into a single system of prevention, treatment, and aftercare? *Issues in Science and Technology* 7:78-84.
- Jones, D. and T. Fisher. 1997. *Community studies report: Little Rock Fighting Back*. Unpublished manuscript.
- Kadushin, C., E. Reber, L. Saxe, and D. Livert. Forthcoming. The substance use system. *The Journal of Substance Use and Misuse*.
- Kaftarian, S. J. and W. B. Hansen. 1994. Improving methodologies for the evaluation of community-based substance abuse prevention programs. *Journal of Community Psychology*, CSAP Special Issue, 3-6.
- Kelling, G. L., T. Pate, D. Dieckman, and C. E. Brown. 1974. *The Kansas City Preventive Patrol Experiment*. Washington, DC: The Police Foundation.
- Kreft, I.G.G. 1998. An illustration of item homogeneity scaling and multilevel analysis techniques in the evaluation of drug prevention programs. *Evaluation Review* 22 (1): 47-75.
- Lorion, L. P. 1994. Evaluating the community partnership program: Reflections on a name. *Journal of Community Psychology*, Monograph Series CSAP Special Issue 201-205.
- Murray, D. M., J. M. Moskowitz, and C. W. Dent. 1996. Design and analysis issues in community-based drug abuse prevention. *American Behavioral Scientist* 39:853-67.
- Parker, E. B., D. T. Campbell, T. D. Cook, N. Katzman, and M. Butler-Paisley. 1971. Time-series analysis of effect of television on library circulation in Illinois. Unpublished paper, Northwestern University.
- Rindskopf, D., M. Stirratt, D. Livert, and L. Saxe. 1997. The utility of multilevel models in the evaluation of community-based programs. Unpublished manuscript.
- Rosenbaum, P. R. 1995. *Observational studies*. New York: Springer-Verlag.
- Saxe, L. and M. Fine. 1981. *Social experiments: Methods for design and evaluation*. Beverly Hills, CA: Sage.
- Saxe, L., E. Reber, D. Hallfors, C. Kadushin, D. Jones, D. Rindskopf, and A. Beveridge. 1997. Think globally, act locally: Assessing the impact of community-based substance abuse prevention. *Evaluation and Program Planning* 20 (3): 357-66.
- Spickard, A., G. Dixon, and F. Sarver. 1994. Fighting Back against America's public health enemy number one. *Bulletin of the New York Academy of Medicine: A Journal of Urban Health* 71:111-35.
- Susser, M. 1995. The tribulations of trials: Interventions in communities. *American Journal of Public Health* 85:156-8.
- Tukey, J. W. 1977. Some thoughts on clinical trials, especially problems of multiplicity. *Science* 198 (4318): 679-84.
- Weiss, C. H. 1980. Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion and Utilization* 1:391-404.
- . 1995. Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In *New approaches to evaluating community initiatives*, edited by J. P. Connell, A. C. Kubisch, L. B. Schorr, and C. H. Weiss. Washington, DC: The Aspen Institute.

- Wiener, N. 1948. *Cybernetics, or control and communication in the animal and the machine*. New York: John Wiley.
- Winick, C. and M. J. Larson. 1997. Community action programs for substance abuse. In *Substance abuse: A comprehensive textbook*. 3rd ed., edited by R. B. Millman and J. G. Langrod. Baltimore, MD: Williams and Wilkins.
- Yin, R. K. and S. J. Kaftarian. 1997. Introduction: Challenges of community-based program outcome evaluations. *Evaluation and Program Planning* 20 (3): 293-8.
- Zeeman, E. C. 1977. *Catastrophe theory: Collected papers 1972-1977*. Reading, MA: Addison-Wesley.

David Rindskopf is a professor of educational psychology and psychology at the City University of New York Graduate School and University Center. He specializes in applied statistics and nonexperimental research design.

Leonard Saxe is a professor of psychology at the City University of New York Graduate School and University Center and a research professor at Brandeis University, the Heller School. He specializes in evaluation of social policy and the social psychology of lying.

Within the context of a large-scale, comprehensive evaluation of the California Drug Alcohol Tobacco Education (DATE) program, this study sought to extend knowledge of student perceptions of prevention education using a naturalistic approach. The constant comparative method was used to analyze 40 focus group interviews of risk and thriving groups conducted in 11 high, middle, and elementary school districts. This article presents three assertions generated solely from 490 "narrative stories" found in the data set. "At-risk" and "thriving" students at all three levels of schooling (a) use "story" to make sense of prevention education, and (b) distinguish use from abuse. High school students of both groups (c) believe that hearing only one side of the substance use/abuse story and strict expulsion policies further alienate students most in need of help. Implications for the use of story as an assessment tool are discussed, as are implications for substance use prevention policy.

THE OTHER SIDE OF THE STORY

Student Narratives on the California Drug, Alcohol, and Tobacco Education Programs

MARIANNE D'EMIDIO-CASTON

University of California, Santa Barbara

JOEL H. BROWN

Center for Educational Research and Development, Berkeley, California

When social problems capture public imagination, public schools often become the vehicle for social change (Sarason 1982). With the focus of public attention, efforts are made by politicians to "do something" to solve the problem. The "something" often results in a mandate to public schools. Directly tied to and intricately connected with state and federal government through funding, schools are readily available settings for delivery of the mandated solution. In the last decade, public schools have increasingly been held accountable to teach young people about the dangers of drugs, alcohol

AUTHORS' NOTE: *This research was supported by the California State Department of Education, Contract No. 3279. It was originally presented at the Annual Meeting of the American Educational Research Association (AERA), April 1995. The views expressed herein are those of the authors and do not necessarily represent those of the California State Department of Education.*

EVALUATION REVIEW, Vol. 22 No. 1, February 1998 95-117

© 1998 Sage Publications, Inc.

and tobacco, AIDS, gangs, and violence. Each of these issues now claims time in the curriculum delivered to students alongside a basic program of academic study. Beginning in the 1980s, "Just Say No!" became the slogan for an ongoing social change effort known as the War on Drugs, aimed at the target of a drug-, alcohol-, and tobacco-free society.

In 1991, the California Drug, Alcohol, and Tobacco Education (DATE) Program was initiated in an effort to consolidate programs to prevent substance use and abuse by children and adolescents. School districts were mandated to provide comprehensive drug, alcohol, and tobacco education for students K through 12th grade. A large-scale evaluation of DATE by the Southwest Regional Laboratory (SWRL) suggests that "at a minimum California schools spent \$83.78 per student in 1992-93 to provide students with prevention education curricula, and positive alternative activities, provide personnel with staff development and Alcohol Tobacco or Other Drugs (ATOD) training in curricula, identification and referral services" (Romero et al. 1994, 38). Since 1991, the cost of DATE has been estimated at \$1.6 billion (Brown, D'Emidio-Caston, and Pollard 1997). Such public focus and fiscal priority on a perceived social problem requires comprehensive evaluation and public accountability.

From 1991 to 1994, an evaluation was conducted along three quantitative dimensions: cost, program implementation, and self-reported student substance use knowledge, attitudes, beliefs, and behaviors (Romero et al. 1993, 1994). Another study, using the same school districts included in the Romero evaluation, looked at the social processes of DATE program implementation (Brown et al. 1993, 1995). These two studies present findings that are often contradictory. Romero, for example, leaves the reader with a positive impression of the effects of DATE. Brown and colleagues are not so convinced of the benefits in relation to students who are labeled most at risk for substance abuse. Although both studies are valuable, an explanation for the discrepancy in findings may be that the voices of students are more clearly heard in the Brown et al. study.

DATE programs have been designed and implemented from a "risk orientation" toward prevention (Brown et al. 1993). A risk orientation includes three characteristics. First, the terms substance use and substance abuse are interchangeable. Second, a risk orientation assumes that a majority of children fall into the "at-risk" category. Thus, "at-risk" is not differentiated from "high-risk." Third, with a risk orientation, there is an absence of focus on resilience as a prevention strategy. As it applies to prevention, the risk orientation is an operational definition of a deficit model, where young people are seen as problems to be fixed rather than resources who make contributions

to their families, schools, and communities (Blue-Swadener 1995; Benard 1993).

There are serious problems associated with the risk-orientation. Using it to inform the solution of one or another perceived social problems masks the underlying social, economic, or environmental conditions that contribute to alienation and hopelessness. Another problem with the risk orientation we argue is that by using broadly defined categories (Hawkins et al. 1987), risk-oriented programs cannot be sufficiently targeted to students most in need of help. Given these limitations, we contend that the risk orientation limits practices in such prevention programs to those that are primarily symbolic (Brown et al. 1997). The appearance of a uniformed officer in classrooms, as in the widely implemented D.A.R.E. (Drug Abuse Resistance Education) program, or Red Ribbons tied around trees during Red Ribbon Week are public displays of something being done about the drug problem. The risk orientation makes it easier to believe that such symbolic public displays are effective programs (Brown and D'Emidio-Caston 1995; Brown et al. 1997).

It is time to sort out the symbolic from the actual effects of DATE services. The story of DATE from the students' point of view is essential to the comprehensive assessment of DATE. If the alienation and hopelessness that young people feel leads to drug or other substance abuse, it is crucial to know whether the risk orientation that guides program development contributes to reducing the alienation and hopelessness. The primary focus of this article is to illuminate the influence DATE services have on students, through analysis of the unsolicited stories students told.

Results of the 1992-1993 qualitative evaluation are reported in Brown et al. (1993). Student perspectives were not included in the first-year results. The second year (1993-1994) qualitative study recognizes the centrality of the learner in prevention education. What meanings do students make of the programs in which they participate? This article posits that the voices of students can be heard through their narrative attempts to make meaning of prevention education. From the story data, we have a better understanding of the answer to the question, Do students perceive that prevention education makes a positive difference in their lives, or are the effects of these programs primarily symbolic?

THEORETICAL FRAMEWORK

"Narrative" is becoming more widely accepted as "a way of knowing" in educational research (Schubert and Ayers 1992; Witherell and Noddings

1991; Connelly and Clandinin 1990; Polkinghorne 1988; Rosen 1985; Mitchell 1980). Mitchell's *On Narrative* brought the study of the role of narrative "out of the realm of the aesthetic into the realm of social and psychological formations," particularly in structures of value and cognition. The study of narrative "has now become a positive source of insight for all the branches of human and natural science" (Mitchell 1980, ix). Cognitive psychologists have been interested in the study of the general structure and function of narrative (Chomsky 1966; Rosen 1985) and the acquisition of narrative skills by children (Bruner 1990; Kemper 1984). The role of narrative in curriculum studies has influenced a reconceptualization of curriculum at the macro and micro levels. Curriculum developed from one perspective has been reconceptualized as the collective story made of multiple perspectives. In "narrative inquiry" (Connelly and Clandinin 1990), researchers seek to understand the ways in which curriculum is constituted in the subjectivity of teachers and other curriculum workers by privileging individual storytelling.

Nevertheless, the role of narrative in evaluation research is in its infancy. Researchers could uncover no work in which narrative as an evaluation tool was applied to substance use and abuse prevention. Researchers did, however, uncover one scientifically sound and germane narrative evaluation. In the "Voices From the Inside" Report (Poplin and Weeres 1992), a "bottom up" narrative approach was taken to examine the state of public schools. Here, they used context-dependent units to produce an infrastructure that, when compared with the primary target population, explains program effects (Patton 1990; Manning and Cullum-Swan 1994). "Voices From the Inside" established the narrative of the target population (presumably the "bottom" in a bottom-up evaluation) in comparison with a given context as an important way to determine program effectiveness. In the Claremont "Voices" study, Poplin and Weeres interviewed teachers, custodians, parents, day-care workers, security guards, cafeteria workers, nurses, and administrators to create a contextual infrastructure for "multiethnic student voices," who formed the centerpiece of their evaluation. By contrasting these contextual voices with the students' voices, they determined that "heretofore identified problems of schooling (lowered achievement, high dropout rates and problems in the teaching profession) are rather consequences of much deeper and more fundamental problems" (p. 11). In both methodology and findings, the "Voices" evaluation represents an important advance in evaluation research.

In conjunction with other methods, in the DATE evaluation, researchers also use the bottom-up narrative evaluation format to help determine program

effectiveness. By interviewing nearly 400 educators, administrators, and community members, the Brown and D'Emidio-Caston (1995) publication described the contextual infrastructure of DATE, contrasting it with the student voices. This research showed that 42.5% of 40 student focus groups (Grades 5-12) reported receiving health/science courses delivered by teachers and 95% of student focus groups reported receiving prevention education from specialists such as D.A.R.E. officers. It was also reported that in delivering prevention, the risk orientation as described above was the dominant context. In this article, with primary focus on prevention-related stories, students are once again the evaluation centerpiece.

Because drug, alcohol, and tobacco education is primarily an effort to influence the knowledge, value orientation, and behavior of students, attention to the construction of meanings revealed through narrative "story" is an exciting and valuable approach. Through the methods of narrative inquiry, our data reveal the construction of students' understanding of DATE. In effect, what they tell us is their side of the story of the War on Drugs.

EVALUATION QUESTIONS

Our evaluation research questions focus on qualitative process and outcome examinations as described by Donabedian (1980), who viewed process as the set of activities that go on within and between practitioners (and in this case service recipients) and the outcome as a change in a service recipient's current and future status that can be attributed to antecedent practices.

This article is based on the assumption that if school programs are effective, such effects will be born out in extensive student interview data regarding program process (how children construct understanding of the effects of substance use) and outcomes (how students feel these understandings have affected their current status as related to prevention programs).

These specific questions focus attention on the process and product of students' meaning making:

- **Process:** In the context of focus groups, how do students at different school levels (elementary, middle, high school) share their understanding of the effects of substance use?
- **Outcome:** How do students perceive the effects of substance use prevention programs?

METHODS

DATA COLLECTION

The 11 of 12 California districts represented in the second-year follow-up evaluation study of DATE were purposely chosen based on the 1992 evaluation of 50 California districts (Brown et al. 1993; Brown, D'Emidio-Caston, and Pollard, 1997). A balance was sought among districts with respect to socioeconomic status (SES), demographics, and average daily attendance (ADA). Of these, 7 were from Southern and Central California, 2 were from the Bay Area, and 2 were from extreme Northern California. One of the state's largest districts was purposefully selected corresponding to the Romero et al. (1994) study. Two schools from each district were randomly assigned by computer selection. In the largest district only, three were selected. Because a detailed description of the methods used to determine participation in this study has already been presented elsewhere (Brown, D'Emidio-Caston, and Pollard 1997), methods are presented here in only as much detail as necessary to provide the reader with an understanding of the analyzed data subset.

From 23 randomly selected schools, two focus groups of students from each school were interviewed. The two groups were chosen by their principal or other delegated authority on the basis of perceived characteristics of "at risk for substance abuse" or "thriving." Criteria for selection for each group were found to be consistent with expectations. For example, inclusion criteria for the perceived at-risk students were the risk factors of low academic achievement and low commitment to school. Criteria for inclusion in the perceived thriving group were characterized by leadership in the school community. The sampling process yielded 40 useable focus group interviews: 20 elementary school interviews, 9 middle school interviews, and 11 high school interviews, representing approximately 240 students. This process generated 18 complete pairs (thriving and at-risk), 3 mixed groups, and 1 unpaired thriving and 1 blank interview due to audio tape malfunction. The three mixed interviews, from the largest school district, offered a means of comparing mixed groups with the risk versus thriving groups. The data presented here are representative of the entire sample of thriving and at-risk student groups ($N = 40$). The student focus groups allowed researchers to evaluate DATE programs from the student point of view.

Students were interviewed by four trained interviewers in focus groups using a semistructured, open-ended interview schedule (Brown et al. 1995). Interviews were subsequently transcribed for analysis.

DATA ANALYSIS

Using the grounded theoretical approach (Strauss and Corbin 1990), conceptual categories were developed inductively from the data and systematically related to one another. Among the categories emerging from the data set was a surprising number of unsolicited stories students used to explain or elaborate their ideas, to give examples of what they meant, or to demonstrate their immediate engagement with the content of the interview question. Stories are distinctive from other interview data in that they illuminate the connections students make of the stimulus topic to what they know in an authentic and recognizable discourse form. Restricting the data analysis to the stories students told increased the internal validity of the data (Goetz and LeCompte 1984).

What "counted" as story? Stein and Policastro (1984) in their studies of what counts as "story" found that no one single structural definition can account for the wide range of compositions people accept as stories. Their work showed that "segments must include at least an animate protagonist and some type of causal sequence before they will be considered a story" (Polkinghorne 1988, 111). Susan Kemper described the simplest form of story as a dyadic event where something happens and the protagonist responds. A more complete definition of the prototypical story identifies a protagonist and a predicament and attempts to resolve the predicament, the outcomes of such attempts, the reactions of the protagonists to the situation, and the causal relationships among each of the elements in the story (Polkinghorne 1988). Many student stories fit Polkinghorne's prototypical story, including all of the required elements. In our analysis, a student statement was considered a story if it had at least one of these characteristics:

1. The statement included at least the elements of a subject and an action related to the use or abuse of substances. For example, R (Respondent): My grandmother is not very old, she's in her 50s and she drinks one beer a day.
2. It was an expression of personal experience or a tale that had been told and passed along to the speaker. For example, R: Deputy J. told us that this one lady sold her baby for crack.
3. The story had a subject who had performed some action or been involved in some event. For example,

R: A lot of people who get drunk and stuff and they go out and do something like usually they'll get in accidents or what happened was I had an uncle, I don't remember his name, but—no Uncle Jack, I think his name was Jack. He was drunk and he went fishing and he was fooling around with the fishing and so he got the hook caught in his leg [several voices: ugh] and so he got gangrene.

R: He got what?

R: Gangrene and died.

Well-formed stories (Burke, in Bruner 1990, 50) include the following five elements: actor, action, goal, scene, instrument. Bruner asserts that when there is a disunity between any of the five elements (trouble), the narrative agent uses the pattern of discourse known as narrative to make sense of the trouble. The addition of "trouble," or what Bruner refers to as a "deviation from canonical culture," provides the stimulus for the telling of the tale. Using Burke's dramatism model, the fishing story is the student's attempt to illustrate his statement that "a lot of people who get drunk go out and get in accidents" (deviation from the canonical culture): Uncle Jack (actor), was drunk (trouble), went fishing (scene), was fooling around (action), got the hook stuck in his leg (instrument), got gangrene and died (goal). This exemplar has all of the required elements of a well-formed story. It is, for the student, a schema for making meaning of the concept "getting drunk." An important caveat for those who find little credibility in the story told above is that the veracity of the story is not as important as the student's use of narrative as a form of communicating his understanding of the concept of getting drunk. Regardless of the truth of the story, it is a recognizable discourse unit that we believe illuminates how this student thinks about the concept.

When encountered in the evaluation of substance use prevention education, the students' stories become an authentic assessment tool to illuminate what young people know, believe, and hope. It is through their stories that students tell us how they connect with their world, how they see themselves as members of school communities, and how they see themselves in relation to the use of substances.

FINDINGS

The results of analysis show that in 40 interviews, there were a total of 494 stories told by students. The stories weave together numerous topics including how students understand the no-use message of DATE, the difference between what they hear in school and what they see at home, their understanding of addiction and of harmful consequences to their health, their understanding of what happens if they get caught using a substance at school, their fears for friends who have substance-related problems, who they think are helping them and who they think are not, and what they think would make a difference. The findings presented in this article, stated as assertions,

illuminate the relationships of the various topics, the process of making sense of prevention education, and the outcomes. The findings are organized in the following manner: First, evidence is presented to support the assertion that students at all three levels of schooling use personal narrative to make sense of the information they receive in substance prevention programs. This assertion corresponds to the evaluation question focused on the process of students' meaning making. Second, building on the evidence presented to support the first assertion, evidence is presented to support the related assertion that by connecting and contrasting the information they learn in school about substance use/abuse with their own experience, students at all three levels in contrast to prevention education programs distinguish use from abuse. This assertion corresponds to both the process and outcome questions. Finally, story evidence is presented to support the assertion that the application of sanctions (detention, suspension, and expulsion) provokes further alienation and disconnection of those students who already see themselves on the periphery of the school community. This assertion corresponds to the outcome question guiding this study. The excerpts provided in all cases represent the predominant point of view found throughout the story data. The excerpts chosen are the most articulate exemplars.

PROCESS

In the context of focus groups, how do students at different school levels share their understanding of what they know about the effects of substance use? Students at all three levels of schooling use personal narrative to make sense of the information they receive in drug prevention education. In the following excerpt, a high school student tells his own story about experimentation with marijuana. He contrasts what he has heard in school with his own experience.

Personal Experience in Narrative Form

R: People say you use it once you're gonna get addicted! I don't see that! But, there, I don't even see, some people say that the drug is addictive, like with a little pressure that you could do anything to keep on using it! Any drug is addictive! And I mean, I, myself, I have smoked marijuana before and I believe it's all in the way you look at it.

I: Uh huh.

R: I tried it and it wasn't nothing, there wasn't anything there for me! People say oh, it makes you feel better and all this stuff, I didn't, there was nothing there for me! And I made my choice to say there's what I thought to myself, what

compels people to do this? Because there was nothing there for me, and I was thinking what is there for them? (0211, ST.S 593, p. 19)

The preceding excerpt is an example of personal experience in the form of narrative story. It offers insights to the meaning the student makes of the prevention education he received. He has clearly not been convinced to forgo experimentation with marijuana. Rather, the information he received conflicts with what others have said, causing disequilibrium, which in turn has prompted personal experimentation. He is struggling to understand the different choices people make. In the next excerpt during a discussion of the various effects of alcohol on people the students knew, two elementary students were moved to tell their own stories:

R: [first] Like, see my uncle, he can drink and he won't get drunk and then my other uncle he can drink a couple of beers and he will get drunk and get into stuff.

R: [second] Like my dad he can drink like three or four beers and he doesn't really get drunk, he gets kind of weird [said with a kind of laugh], but he doesn't get drunk and if my mom if she drinks anything alcoholic she gets really sick, because he, I mean, my dad used to drink more than he does now. I mean, lately he has maybe one beer a month and my mom doesn't drink. So, it just kind of depends on the attitude of the person they drink, too, because if they're already violent then if they drink they might get even more violent and then if it doesn't bother them, you know.

I: Does the D.A.R.E. officer teach you those things?

R: [third] No, not really.

R: [different respondent] I don't think so.

I: So how did you come to know that? Just by watching?

R: [second] You just kind of know it. [short laugh] [second respondent says "Yes" in the background] You know just by observing your surroundings and you can tell how people act. I mean, all families have different examples of stuff but you can just about get in any family somebody that drinks. (0027, ST.E 567 p. 6-7)

The stories told by these elementary students are typical of both risk and thriving groups. They are aware that alcohol has a negative effect on the behavior of some people. They are also aware that others who drink do not have a problem and can use alcohol occasionally. Most important, the D.A.R.E. officer has not given them this message. As the high school excerpt also illuminated, they have constructed it from their own personal observations and experiences. Through the stories about uncles' and parents' alcohol use, they reveal the understanding they have of use ("he can drink and won't

get drunk") and abuse ("he can drink a couple of beers and he will get drunk and get into stuff").

In the following middle school excerpt, during a discussion of what the students think should be in the curriculum they receive, the researcher asked a question that prompted the student to talk about her parent's enjoyment of wine.

I: Well then, what would you guys like to see in the classes that you don't get now?

R: Two sides of the story.

R: Yeah, we . . .

I: Wait. Can you explain to me what you mean by two sides to the story?

R: Because they give one side, telling you how bad it is, and then they should have another side saying, cause well, they always tell us drinking it really bad and don't drink cause you get drunk and you end up killing people and yourself. But, that's not true cause they tell you that one glass of wine could do that! But, I think they all had another side. That it's okay if you have a little, but not get drunk.

R: Yeah, because everybody is going to drink when they get older! Maybe just, I mean my parents enjoy a glass of wine with dinner and that's just the way it goes! [laughs] It's not like we can stop them from having a . . .

I: Well, would you want to stop them from having a glass of wine with dinner?

R: No, because I think they enjoy it. They don't get drunk on one glass of wine! [laughs] I think they enjoy having a glass of wine once in a while. They go up to Napa and get some nice aged wine and have some nice wine with dinner or at a party. I wouldn't want to stop them from doing something that they enjoy! (0005, ST.M 507, p. 19)

One of California's largest industries is wine making, as many California students are aware. By telling the story of her parents' trip to Napa, this student demonstrates an awareness of the culture that enjoys wine with dinner. She believes that everyone will drink when they get older. She is also able to distinguish use (enjoy a glass of wine with dinner) from abuse (its OK if you have a little but not get drunk), and she is outspoken in her desire to hear both sides of the story from those who deliver substance use prevention education. More significant, she says clearly that what she has heard in school is not true.

The excerpts presented above are representative of 38 of 40 interviews. In each case, the student uses his or her own personal experience or a significant other's experience with a substance to make a connection with the information he or she has received in school. It is apparent from the above excerpts that the students use narrative to not only link their personal experience to what they have learned in school but also to contrast it.

Close examination of the three excerpts above reveals that in each case, what students learned in their substance abuse prevention education is not consistent with other life experiences. In the elementary school excerpt, two students present stories. The first story is about an uncle who is harmed by using alcohol and an uncle who is not. The second student describes the different reactions of his mother and father to the use of alcohol. These students are aware that different people have different reactions to the use of alcohol. The D.A.R.E. officer has not given them this information; they have constructed it from their observations of people in their lives. In the middle school story, the student contrasts her parents' enjoyment of wine with the no-use message she has heard at school. The distinction is not present in the education she receives, and she is clearly aware of the difference labeling the prevention message "untrue." In the high school excerpt, the student contrasts his own experience of using marijuana with the two different ideas he has heard about the use of the substance. He has been taught that "if you use substances you will get addicted." Others in his experience have told him it "will make you feel better." His personal experimentation has not confirmed either of the two predictions. Bruner's assertion that stories are stimulated by the mismatch of an event and the "canonical" would certainly seem to be operating here.

All of the preceding leads to a more developed version of the process assertion. Through the narrative form, students in our study relate the experiences they have in their personal lives to the information they receive at school. By linking and contrasting the two experiences, they construct their own understanding of the effect of using drugs, alcohol, and tobacco.

OUTCOMES

How do students perceive the effects of substance use prevention programs? In the next section, we will make more explicit the contrasts between prevention education and the students' constructed understandings. When students contrast their experiences with what they are taught, a common theme emerges. The theme corresponds to the outcome question guiding this evaluation. Students perceive the effects of prevention education as having little influence on their decision making. We have presented evidence throughout the article to support the assertion that students construct their own understandings of the effects of the use of substances. When students' understandings are different enough from the message they receive in DATE,

the credibility of the information they receive and in some cases the students' trust in those who offer the information may be called into question. The following example illustrates a student's blatant distrust of the information he has received.

Student Constructed Understandings

R: No, I don't believe that stuff about one cigarette! No! My mom smokes to calm her, my mom is really hyper person and she smokes to calm her nerves. She's allowed to do it, she works, she pays her bills, so she's allowed to do it! (0005, ST.M 507, p. 23)

The middle school student is certain about his mother's right to smoke when she wants to calm her nerves. He appreciates the fact that she is a responsible adult and can make her own decisions.

Students at all three school levels are able to distinguish use from abuse. The following representative excerpts at each of the three school levels constitute evidence that students distinguish between use and abuse. The stories students told distinguishing use from abuse often included their personal experiences. Although elementary students are legally prohibited from drinking, it must be acknowledged that many of the elementary students have tried alcohol in one form or another under various conditions.

Distinction of Use and Abuse

R: But if you drink like too much alcohol at one time, too fast, it happened to me once, it was an occasion and I had a little shot of wine and I was thirsty and I drank it all at once because I was really thirsty and five minutes later I was sort of snoring.

I: [laughs] Right, right.

R: I'm in the seat going [makes snoring noises].

I: So you were out, huh?

R: Yes. I'm not going to do that again. (0072, ST.EH 533 p. 8)

Students at the elementary level are aware that drinking "too much" "too fast" is abuse. The story illustrates the power of personal experience to teach and reinforce lessons that adults would like students to learn. The middle school students in the following excerpt use story to support their conclusion that not everyone who tries alcohol or drugs "has a problem."

R: I have a friend in high school and she used to do alcohol and she quit. She used to do drugs, but she quit. It's very easy to quit! If you put your mind to it.

R: Some people it's easy for, some people it ain't.

I: So, you don't think that everybody that tries it has a problem?

R: Right. No. (0005, ST.MH 508, P. 8)

Stories told throughout the data illuminate students' understanding of what constitutes a "problem" or "abuse" of a substance. "Being able to stop" is one way students identify who does and who does not have a problem. In the next excerpt, a story of a person who "can't stop" offers an example of what the high school student sees as the road to alcoholism.

R: My friend's girl has 3 or 4 beers and she'll get real buzzed and she has to keep drinking more and more! She can't just enjoy it, she has to get loaded. She can't stop! I can just walk away from it anytime, or drink several and have a buzz and be alright. And I see her, I don't like people who drink to get drunk! You know, just to drink?

R: Yeah!

R: People like that are turning into alcoholics! You can see it coming! (0185, ST.SH 545, p. 14)

The stories selected to support the notion that students distinguish between use and abuse are, again, typical of those found in the majority of interviews. The importance of this assertion is understood in the context of the clear message presented to the students at all grade levels that use of substances equals abuse. Students typically understand that all use of alcohol is not abuse, and they clearly identify what is abuse. The disparity between what they are taught and what they present as story demonstrates that the no-use message is not being "taken up" (Bruner 1990, 63).

From the excerpt presented above, an extension of the disparity between what is taught in school and what is understood by students is uncovered. Many students not only differentiate use from abuse, they believe that a person has to want to stop abusing substances for counseling or sanctions to have an effect. The idea that it is easy to quit for some people and more difficult for others is linked to a story about a high school friend who was successful when she "put her mind to it." This story illuminates an important issue for students at all grade levels but most notably at the high school level when young people are most likely to start using substances. Students believe that it is up to the person to want to stop. Neither counseling nor sanctions levied against students who are caught using have much preventive influence.

R: Um, no. I don't think that counseling can really, it can help you, but I don't think it's gonna change your mind. You have to be willing to change your mind! To not to do it, or to not want do it. If you go to counseling and they tell you it's all bad, but you still think it's good, then you're gonna do it! (0005, ST.SH, p. 12)

This extension of the assertion will also be discussed.

Inconsistent Message of Home and School

If what is taught in school is not being accepted by students, is it because what they learn in school is different from what their parents say and do? As in several of the prior story examples, dissonance occurs when students witness their parents' use of substances. They are forced to deny what they learn in school, "I don't believe them . . ." or make a judgment about their parents, "she had a right to smoke. . . ." The following elementary excerpt is presented in order to make explicit the lack of consistency between what parents are telling their children and what the school is telling them.

I: OK, but what I'm asking you guys—this is a very personal question—what I'm asking you guys is how do you decide that a little bit is OK and a lot is too much? Did someone tell you that?

R: Yes.

I: Or did you just make up your mind on your own?

R: My dad when he was—I don't remember how old—he told me that he was with his friends at a party and they told him to try a beer and so he said OK and so he drank one and then drank another and he started getting sick and he threw up so since he's only drank like a half a beer or something so he doesn't get sick any more.

I: So did most of you get that idea—is he right and most of you got that idea from your parents?

R: Yes. [several voices] (0072 STE 532 p. 19)

The notion that students use narrative story to construct meaning of their diverse experiences with substance use and abuse is very powerful. When there is a mismatch between home and school, the student is forced to resolve the dissonance she or he experiences by making sense of the two worlds. In effect, the students are being asked to make choices between two authorities, both of whom lose credibility in the students' eyes in too many cases. Often, the dissonance results in undermining the students' trust in adults in general.

Undermining of Students' Trust in Adults

Analysis of the stories found in the high school data revealed a general outcome related to the lack of trust, but having even more serious consequences in the "high risk" population (Hawkins et al. 1987) the very student prevention programs were originally intended to help. The following excerpt is typical of students who see themselves outside the school community.

R: I mean they always do it like we're all bad people here.

R: I don't think the schools are for like helping it's just for getting the bad kids out and it's just . . .

R: Yeah.

R: Well, maybe if you could get them to care more then they would do that [a different respondent than the others above]

R: If they suspect you of smoking or having drugs on you or whatever, if they see a kid like that in their school then, instead of suspending them and getting them out of school, why don't they help them? (0072, ST.SH 531, p. 13)

These at-risk students, according to the "risk factor model" (Hawkins et al. 1987), are the most likely to become dropouts, drug addicts, homeless, or criminals. Yet, all too often these young people feel hopeless and disheartened and see no future for themselves in the school or society. Another excerpt gives further insight to the minimal effects of prevention education.

R: It's pretty sad if society puts you in a position where you can't be happy unless you use drugs. I mean if you got school and you got the wrong problem, not a drug user, but about the way society treats kids. (0072, ST.SH 531, p. 10)

These students believe the treatment (prevention education) is for the wrong problem. They see themselves as victims of social pressures, and they are concerned about the lack of care and support they receive from school personnel to cope with these perceived pressures.

If only the voices of at-risk students were raised urging those in authority to help, they would probably not be heard. However, they are not the only voices urging a change in the way students are treated when they have a problem. School personnel recognize the failure of the school system to help these students as well. "We still get rid of too many kids . . . those are the kids that the state of California and the United States of America have identified as their target population. . . . The kids that are at risk the most, are the kids that are exited from the system and they do not have access to the resources. . . . The kids that we need to help in and provide resources to are the kids that we exit from the system" (0027, GF 558, p. 18).

Given the previous data, we come to the final assertion generated in this study: The application of sanctions (detention, suspension, and expulsion) provokes further alienation and disconnection of those students who already see themselves on the periphery of the school community.

The next section of the article will discuss the implications of the preceding assertions.

DISCUSSION

Returning to the first of the two questions that guided this evaluation, it is apparent from the evidence supporting the assertions that the students of both groups, at all three levels of school, use narrative story to display their understandings of substance use and abuse. The ubiquity of this form of discourse in the student data adequately supports the proposition that story provides a way of sorting out our thoughts about the world. The student stories also support Bruner's idea that narrative mediates between the canonical world of culture and the idiosyncratic world of beliefs, desires, and hopes. If stories are the medium by which human beings construct meaning, we argue that the student stories found in the interviews are a key to understanding how students are making sense of the programs they receive. Unsolicited stories were woven throughout all but two of the interviews. Curiously, these two interviews were conducted by the same interviewer whose style of interaction with the students included interrupting them while they were speaking and making references to time during the interview. This interview style undoubtedly contributed to the lack of stories. Excerpts from the interviews have adequately shown how the students, stimulated by the conversation, voluntarily share the stories they associate with the stimulus. This primary assertion supports Polkinghorne's (1988) notion that "experience is constructed when a person assimilates the stimuli and matches them with his or her existing structural representations of events which are judged to be similar to the input given" (p. 108). During the interviews, questions were asked that stimulated the mental representations of similar events (stories) that, in the student's mind, matched the stimuli.

In analyzing the data, we did not view the stories of students uncritically. The DATE evaluation used multiple methods to assess program effectiveness, and narrative story was one of them. Narrative stories were not anticipated in our data collection process. It was the overwhelming number of stories that the students told that focused our attention on the value of narrative. Our primary concern is not the factual basis of these stories. As we have shown,

whereas many stories may represent facts, others represent misconceptions or partial truths regarding substances like alcohol. We see, through students' stories, as in the construction of understanding of other types of knowledge, the logic the student uses to make sense of the world. The importance of narrative as an evaluation tool is twofold. It features the voices of the target population at the center of the evaluation of programs, and it helps evaluators gain insight to the construction of meaning students are making. In this study, when students told their stories, we gained insight to what they have learned and how they make sense of prevention education. Viewing these findings critically, we feel reassured by the triangulation of other results from different data sources in the DATE evaluation (Brown, D'Emidio-Caston, and Pollard 1997; Brown and D'Emidio-Caston 1995).

Regarding the outcomes of prevention education and what we now understand as the mismatch between prevention education and personal experience, we can begin to sort out the effects. In some cases, the stories told were simple accounts of someone's use of a substance. In other cases, they are elaborate, well-formed stories that illustrate the students' confusion, disequilibrium, or dissatisfaction with the lack of consistency between their personal experience and what the school authorities tell them. Students' ability to distinguish between the use and abuse of substances is an indicator of such lack of consistency. The narrative evidence revealed how the students interpret and connect what they learn in school with what they experience out of school in the popular culture and home environment. When a student's home life includes drinking wine with dinner, for example, or one parent's capacity to drink and another not, there is a problem with telling that student that all drinking is unhealthy or bad. They must resolve their disequilibrium, and often do, at the expense of not believing the information or the person who delivers the inconsistent message. When that person is a teacher or a police officer in the D.A.R.E. program, the unfortunate result is a loss of credibility in those who represent social authority.

For many students, particularly those who are active, thriving members of the school community, the loss of credible authority in the form of teachers and police officers is not alienating. These students see themselves as members of the school community. They perceive that the reason behind the inconsistent message is good will and "caring" for their well-being. The unfortunate antithesis of this is true for those who are already on the periphery of the school community. For the students who have "low commitment for school," the loss of credible adult authority pushes them further toward the periphery.

Clearly, the hard line policies called for by the DATE application are successful in reducing the number of students with drug-related problems in

the schools. Equally as clear is the unfortunate way this outcome is enacted. The schools do not seem to have the capacity to help or heal. They have only the capacity to punish and expel. Those students who perceive themselves as "bad" have no incentive whatever to comply with the no-use policies (Napier and Gershenfeld 1993). For them, detention, suspension, and expulsion confirm their perceived non-member status. These implications undermine the position that a risk orientation is a valuable tool to change patterns of substance use or abuse in young people. We argue here that it would seem appropriate and propitious to change the assumptions guiding the substances use prevention programs in California public schools.

Others too have urged a different approach. Benard (1993) and Brown and Horowitz (1993) have clearly stated a different orientation to working with students who see themselves as alienated from the school community. Benard urges schools to become places characterized by caring, participation, and high expectations for all students. Her argument is that when students feel connected to the school community, they feel cared for and they have better resiliency and healthy responses to challenges. Brown and Horowitz urge a "harm-reduction" model that reduces the actual damage a person might experience from secondary causes related to use of substances. Designated driver programs are one example of a harm-reduction strategy.

What do students say? It is fitting to end this article with some final excerpts from students who have a great deal more knowledge than we often credit them. When asked what the goal of a drug education program should be, this high school student replied:

To know what your limitations are, to make yourself aware enough so that you know—personally, I've never felt very worried that I would ever become a substance abuser. When I was like elementary school it was crammed down my throat, Just Say No, it's the most awful thing in the world, and so when it first came, like in ninth grade, I remember this girl was trying to get me to do pot I'm like, "No, that's evil." It was that kind of a thing, but I think the goal of education should be you're going to be in the situation, you're going to see this, that and the other thing, it's not evil if you've got a good enough sense of self worth, if you know what your boundaries are, if you know what you feel comfortable with and if you know what it's going to do to you and you know what the consequences may be. (0072, ST.S 530 p. 15)

Her recommendation that students need to have a good sense of self-worth and know what their boundaries are resounds the wisdom of the adults cited above. If the school creates a climate where all students experience success and a sense of accomplishment, they will be more resilient when faced with

the givens of conflicting authorities or economic hardship. Another student had this recommendation:

I just want to say that I guess the best education would be the education that would allow you to evaluate yourself and allow you to evaluate your own personal beliefs and your morals and your values and take a strong look at what you're feeling and if you might have the possibility to be a substance abuser. (0072, ST.S 530 p. 31)

The figure attached to the DATE Program during the years of this evaluation in the state of California is estimated at more than \$1.5 billion. Public accountability for this large an expenditure is appropriate. Our research has shown that risk-oriented policies and programs like D.A.R.E., Red Ribbon Week, and anti-drug assemblies are highly implemented. Their primary program components are some form of scare tactics, offering a reward in exchange for not using substances and enhancing self-esteem through refusal skills. Policies widely in place are intended to enforce the social and legal consequences of substance use (Brown et al. 1997; Brown and D'Emidio-Caston 1995). The stories presented in this article are representative of hundreds of stories the students in the DATE evaluation told. It is clear that they do not believe what they are being told and instead construct their own version of the consequences of substance use. The DATE evidence stands with other evidence in suggesting a high level of program implementation and low level of effectiveness (Klitzner 1987; Moskowitz 1989; Tobler 1992; Ennett et al. 1994). We have presented an argument here that demonstrates that prevention programs designed with the risk orientation have a potentially more insidious effect, that of reinforcing the perception of alienated young people that adult authorities are not credible or caring. We suggest we listen to their voices as they tell us we are treating the wrong problem. In examining and observing programs and program records, performing interviews, doing surveys, and performing meta-analyses of other study results, we are left with few alternative explanations in our inability to show positive program effects.

The War on Drugs has had many casualties. Our results indicate that students who demonstrate the need for the most support may be unintended victims of that war; not from the use of substances themselves but from the process of substance use prevention education and the policies in place in school districts, which exclude them. Those students who are thriving, although they may experiment, have good reason for not abusing substances. They see themselves in the future, and they have legitimate, school-sanctioned support networks. Those who abuse substances are often those with little vision of themselves in the future. Without a legitimate, sanctioned support

system, they may seek in gangs the affiliation and recognition society has withheld. Without condoning the use of substances by young people, a more authentic and realistic orientation to working with students who have problems must be found. Emphasis on resiliency and harm reduction are two possibilities. With each day, as our jails take up more and more of the available resources, an ever greater need is apparent. For prevention programs to be effective, they must support those most at risk to be able to see a future when they close their eyes and dream.

REFERENCES

- Benard, B. 1993. *Turning the corner risk to resiliency*. San Francisco: Western Regional Center.
- Blue-Swadner, B. and S. Lubeck. 1995. *Children and families at promise: Reconstructing the discourse of risk*. New York: SUNY Press.
- Brown, J. H. and M. D'Emidio-Caston. 1995. On becoming at risk through drug education: How symbolic policies and their practices affect students. *Evaluation Review* 19 (4): 451-92.
- Brown, J. H., M. D'Emidio-Caston, T. Goldsworthy-Hanner, and M. Alioto. 1993. *Technical report of 1992 qualitative findings for the drug, alcohol, and tobacco education program evaluation*. Los Alamitos, CA: South West Regional Laboratory.
- Brown, J. H., M. D'Emidio-Caston, K. Kaufman, T. Goldsworthy-Hanner, T. and M. Alioto. 1995. *In their own voices: Students and educators evaluate school based drug alcohol and tobacco education programs*. Prepared for the California Department of Education by Pacific Institute for Research and Evaluation.
- Brown, J. H., M. D'Emidio-Caston, J. A. Pollard. 1997. Students and substances: Social power in drug education. *Educational Evaluation and Policy Analysis* 19 (1): 65-82.
- Brown, J. H. and J. E. Horowitz. 1993. Deviance and deviants: Why adolescent substance use prevention programs do not work. *Evaluation Review* 17 (5): 529-55.
- Bruner, J. 1990. *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Chomsky, N. 1966. *Cartesian linguistics: A chapter in the history of rationalist thought*. New York: Harper and Row.
- Connelly, F. M. and D. J. Clandinin. 1990. Stories of experience and narrative inquiry. *Educational Researcher* 19 (5): 2-14.
- Donabedian, A. 1980. *Explorations in quality assessment and monitoring: Vol. 1. The definition of quality and approaches to its assessment*. Ann Arbor, MI: Health Administration Press.
- Ennett, S. T., N. S. Tobler, C. L. Ringwalt, and R. L. Flewelling. 1994. How effective is drug abuse resistance education? A meta-analysis of project D.A.R.E. outcome evaluations. *American Journal of Public Health* 84 (9): 1394-401.
- Goetz, J. P. and M. LeCompte. 1984. *Ethnography and qualitative design in educational research*. Orlando, FL: Academic Press.
- Hawkins, J. D., J. M. Kuagbwem, J. M. Jenson, and R. F. Catalano. 1987. Delinquents and drugs: What the evidence suggests about prevention and treatment programming. In *Youth at high risk for substance abuse* (DHHS Publication No. ADM 87-1537; reprinted 1990 as ADM 90-1537), edited by B. S. Brown and A. R. Mills, 81-131. Washington, DC: U.S. Government Printing Office.

- Kemper, S. 1984. The development of narrative skills: Explanations and entertainments. In *Discourse development: Progress in cognitive development research*, edited by S. A. Kuczaj, 99-124. New York: Springer-Verlag.
- Klitzner, M. D. 1987. *Part 2: An assessment of the research on school-based prevention programs*. Report to Congress and the White House on the nature and effectiveness of federal, state, and local drug prevention/education programs. Washington, DC: U.S. Department of Education.
- Manning, P. K. and B. Cullum-Swan. 1994. Narrative, content, and semiotic analysis. In *Handbook of qualitative research*, edited by N. K. Denzin and Y. S. Lincoln. Newbury Park, CA: Sage.
- Mitchell, W.J.T., ed. 1980. *On narrative*. Chicago: University of Chicago Press.
- Moskowitz, J. M. 1989. The primary prevention of alcohol problems: A critical review of the research literature. *Journal of Studies on Alcohol* 50:54-88.
- Napier, R. W. and M. K. Gershenfeld. 1993. *Groups theory and practice*. Boston: Houghton Mifflin.
- Patton, M. Q. 1990. *Qualitative evaluation and research methods*. 2d ed. Newbury Park, CA: Sage.
- Polkinghorne, D. E. 1988. *Narrative knowing and the human services*. New York: State University of New York Press.
- Poplin, M. and J. Weeres. 1992. *Voices from the inside: A report on school from inside the classroom*. Claremont, CA: The Institute for Education in Transformation at The Claremont Graduate School.
- Romero, F., J. Bailey, C. Carr, J. Flaherty, T. Fleming, J. R. Gaynor, D. Houle, R. Karam, M. Lark, T. Martino, and C. Thomas. 1994. *1992-93 California programs to prevent and reduce drug, alcohol, and tobacco use among in-school youth: Annual evaluation report*. Prepared for the California Department of Education by Southwest Regional Laboratory. Los Alamitos, NM: The Southwest Regional Laboratory.
- Romero, F., C. Carr, J. Pollard, D. Houle, J. Brown, J. R. Gaynor, T. Fleming, J. Flaherty, T. Martino, and R. Karam. 1993. *Drug, alcohol and tobacco education evaluation: Second year interim evaluation report*. Prepared for the California Department of Education by Southwest Regional Library. Los Alamitos, NM: The Southwest Regional Laboratory.
- Rosen, H. 1985. The importance of story. *Language Arts* 63 (3): 226-37.
- Sarason, S. B. 1982. *The culture of the school and the problem of change*. 2d ed. Boston: Allyn & Bacon.
- Schubert, W. H. and W. C. Ayers, eds. 1992. *Teacher lore, learning from our own experience*. New York: Longman.
- Stein, N. L. and M. Policastro. 1984. The concept of story: A comparison between children's and teacher's viewpoints. In *Learning and comprehension of text*, edited by H. Mandl, N. L. Stein, and T. Trabasso, 113-55. Hillsdale, NJ: Lawrence Erlbaum.
- Strauss, A. and J. Corbin. 1990. *Basics for qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Tobler, N. S. 1992. Drug prevention programs can work: Research findings. *Journal of Addictive Diseases* 11 (3): 1-26.
- Witherell, C. and N. Noddings, eds. 1991. *Stories lives tell*. New York: Teachers College Press.

Marianne D'Emidio-Caston is presently the coordinator of the Elementary Teacher Education Program and Lecturer in the Graduate School of Education, University of California at Santa Barbara. She has more than 20 years of field experience as a teacher and administrator and has published articles on the development of practical theory in beginning teachers and the use of the affective domain to enhance teaching and learning.

Joel H. Brown is an educational consultant. Currently, he directs the Center for Educational Research and Development (CERD), a Berkeley-based educational evaluation, research, and policy development organization focusing on the development and well-being of youth. Among many other studies, he was the principal investigator of the nationally recognized research entitled "In Their Own Voices: Students and Educators Evaluate California School-Based Drug, Alcohol, and Tobacco Education (DATE) Programs." His areas of expertise, all within social welfare, public health, and psychological contexts, include the study of adolescent development, organizational change, school evaluation program/policy development, and the integration of qualitative with quantitative research methods.

This article examines developments in school-based drug prevention policy and programming since the Anti-Drug Abuse Act of 1986. Using data from national surveys and evaluations of school-based programs, it argues, first, that there was really no need for a massive infusion of money into school-based drug prevention in the late 1980s, and, second, that there was little or no evidence to indicate that a "new generation" of effective programs, based on the so-called social influence model, was emerging at this time. Despite the infusion of resources into school-based prevention efforts, adolescent drug use has risen in recent years. Moreover, evaluations continue to show that the effectiveness of social influence programs is very much in the eye of the beholder. Fundamental questions need to be asked of school-based drug prevention—just as they should be asked of other key components of our current drug control policy.

THE IRRELEVANCE OF EVIDENCE IN THE DEVELOPMENT OF SCHOOL-BASED DRUG PREVENTION POLICY, 1986-1996

D. M. GORMAN

Rutgers—The State University of New Jersey

It is the declared policy of the United States Government to create a Drug-Free America by 1995.

Anti-Drug Abuse Act of 1988

What is intellectually interesting about visions are their assumptions and their reasoning. But what is socially crucial is the extent to which they are resistant to evidence.

Thomas Sowell (1995)

In his 1995 book entitled *The Vision of the Anointed*, Thomas Sowell proposes that many social policy initiatives of the past 30 years have been

AUTHOR'S NOTE: *I thank Tony Carpenter, Dwight B. Heath, Erich Labouvie, and James Langenbucher for comments on various aspects of this article.*

EVALUATION REVIEW, Vol. 22 No. 1, February 1998 118-146
© 1998 Sage Publications, Inc.

unsuccessful in achieving their stated goals and objectives, but despite this they manage to survive and in many cases thrive under successive government administrations. He argues that there is a characteristic pattern to the evolution of these policies that has four stages: First, a situation is identified and characterized as a "crisis"; second, policies and programs are proposed as a "solution" to this crisis; third, the "problem" that the policies and programs were meant to ameliorate gets worse; and fourth, advocates of the policies and programs develop a response in which it is asserted that without the policies and programs, the situation would be even worse.

Sowell observes that empirical evidence is largely irrelevant to this process—data are used selectively to support the implementation and maintenance of policies and programs, and not to systematically test opposing theories about their effectiveness. Programs can be proven to have "worked" because the standards by which they are judged can be continuously lowered. Questions are raised about those who dispute claims of effectiveness regarding the chosen policies: The issue becomes not whether the stated goals and objectives of programs are being met, but the commitment and motives of critics. The burden of proof is placed on them to demonstrate the detrimental effects of policies and programs, whereas advocates are free to make lavish claims of success and value to society.

The present article uses the framework described by Sowell to examine the evolution of school-based drug prevention policies in the United States during the past 10 years—that is, since the landmark Anti-Drug Abuse Acts of 1986 and 1988, which resulted in a huge expansion of the role of federal government in drug control activities. It looks at the "crisis" of adolescent drug use in the early 1980s that was used to justify this increased role, the rise of school-based drug prevention from the mid-1980s, changes in adolescent drug use in subsequent years, and the response of advocates of school-based drug prevention to these changes. Throughout, attention is paid to the use made of empirical research, especially data generated from evaluation research and large-scale surveys, in the development of school-based drug prevention policy.

THE CRISIS

Trends in adolescent drug use can be tracked in the United States from data collected through two large-scale national surveys—*Monitoring the Future* (Johnson, O'Malley, and Bachman 1996), which reports continuous

data from 1981 onward, and the *National Household Survey* (Substance Abuse and Mental Health Services Administration 1996), which was conducted every 3 years between 1976 and 1988 and annually after 1990. Both surveys show that illicit drug use (composed primarily of marijuana use) among adolescents peaked in 1979, with 39% of 12th graders in the former survey reporting use during the previous 30 days and 18.5% of 12- to 17-year-olds in the latter reporting use during a similar time period. By 1985, 1 year before the first Anti-Drug Abuse Act, the proportion reporting use in each survey had fallen to 30% and 15%, respectively. By 1988, the year of the second Anti-Drug Abuse Act, reported use in these age groups was down to 21% in *Monitoring the Future* and 9% in the *National Household Survey*.

Those behind the buildup of the federal drug control policies were, of course, aware of these data. Explaining the rationale behind the new policies, William J. Bennett in his introduction to the first *National Drug Control Strategy* wrote that although surveys showed that drug use among teenagers was on the decline, available evidence also indicated an increase in "drug-related chaos" such as violent crime and medical emergencies, and that this could be explained by the appearance of crack cocaine in the inner cities. The country was fighting, he added, two drug wars: one "against 'casual' use of drugs by many Americans, and we are winning it," and one "against addiction to cocaine. . . . And on this second front, increasingly located in our cities, we are losing—badly" (Bennett 1989, 4). For Bennett, these two fronts were not unrelated, as addicts inevitably started as casual users. Thus, if the pool of the latter was reduced, addiction rates would eventually also fall. They were also linked at a more fundamental level that had to do with the changing norms and values of American society during the previous two decades, notably, increasingly permissive attitudes toward drug use. For Bennett, the "drug-using flower children of the late 1960s set the stage for the drug gangs of the late 1980s" (Bennett 1992, 123).

In fact, the establishment of the crack economy in many cities in the United States appears to have resulted from forces specific to time and place (Dunlap and Johnson 1992; Hamid 1991), and use of the drug never did spread much beyond poor urban communities (Reinarman and Levine 1995). However, the decision to conceptualize the "drug problem" as universal and a threat to all Americans was an important one, as it clearly influenced the type of policies and programs put in place at the time. Instead of focusing resources and energies on the specific problem at hand (the urban crack economy), a set of broad-based interventions intended to change prevailing attitudes and norms was developed, of which school-based drug prevention was a key component.

THE SOLUTION

As noted above, federal spending on drug control policies increased substantially in the late-1980s, following the legislation of 1986 and 1988. Between 1986 and 1990, the total drug control budget rose from just under \$3.9 billion to \$11 billion.¹ Whereas two thirds of this was allocated to supply-side strategies centered on international interdiction and domestic law enforcement, demand-side activities focused on treatment and prevention also experienced increased funding. The prevention budget increased by close to \$400 million as a result of the first Anti-Drug Abuse Act—rising from \$195 million in 1986 to \$577 million in 1987. Following the Anti-Drug Abuse Act of 1988, the budget rose to \$870 million, and by 1992 the federal government was spending almost \$1.7 billion on drug abuse prevention efforts.

Among the new initiatives established under the 1986 legislation was the Department of Education's Drug-Free Schools and Communities program, the purpose of which was to establish drug abuse prevention and education programs in schools through the provision of federal financial assistance. School-based drug education and prevention, it was stated, were "essential components of a comprehensive strategy to reduce demand for and use of drugs" (Anti-Drug Abuse Act of 1986, sec. 4102). They were necessary, it was argued, as "drug use and abuse are widespread among the Nation's students, not only in secondary schools, but increasingly in elementary schools as well" (sec. 4102). The Department of Education drug prevention budget underwent a massive increase between 1986 and 1987—rising from just \$3.9 million to \$263.9 million. By 1992, the budget was more than \$660 million.

We now tend to take the existence of drug prevention in schools as given, and any suggestion that funding of such activities cease elicits opposition from all parts of the political spectrum (Gorman 1997). However, if one considers the evaluation findings available to policy makers in the mid-1980s, the decision to pursue this approach was by no means obvious.

PROGRAM EVALUATIONS: PRE-1986

Research into the effects of school-based drug prevention programs began to appear in the United States in the late 1960s. Almost without exception, these studies were methodologically weak evaluations of primarily knowledge-based programs. In an early review, Braucht et al. (1973) concluded "that there is almost no empirical evidence of the effectiveness of these programs" (p. 1279). Three years later, Randall and Wong (1976) and Berberian et al.

(1976) reached much the same conclusion. Both of these reviews drew attention to the methodological weaknesses of existing evaluations, the paucity of data indicating effects on drug use behaviors, and the existence of data indicating that drug education efforts might actually be counterproductive. Kinder, Pape, and Walfish (1980) reviewed evaluation studies from the late 1960s and early 1970s, most of which were concerned with information-based programs. As with earlier reviews, they concluded that these programs were ineffective in reducing drug use and might even serve to exacerbate the problem. Goodstadt (1980), addressing the issue of counterproductivity in greater detail, concluded that the available evidence indicated "that 'negative' program effects were not an isolated phenomena, but occur frequently enough and affect self-reported behavior often enough to require more careful scrutiny" (p. 94).

Thus, by 1980, there was little evidence available from program evaluations to support the idea that school-based education was among the "essential components" of a comprehensive drug control strategy. Indeed, in the opinion of many researchers, such education was apt to do more harm than good. As I have noted elsewhere, this pessimism began to be displaced in the early 1980s by the view that effective school-based drug education could be developed (Gorman 1997). One of the earliest research reports to express such optimism was Schaps et al.'s (1981) review of 127 program evaluations. Although concluding that the majority of these produced "only minor effects on drug use behaviors," analysis of a subgroup of 10 exemplary studies led the authors to be encouraged about the efficacy of a "new generation" of prevention programs (although their analysis did not allow description of the common components of these). During the next 5 years, however, the argument began to be developed that successful programs came in one of two basic forms—those focused specifically on drug resistance skills (resistance skills training; RST), and those broadly focused on enhancing general life skills (social skills training; SST). Starting in the late 1970s, evaluations of the effects of these so-called social influence programs on cigarette smoking began to appear, and the National Institute on Drug Abuse published a monograph reviewing these studies in 1985 (Bell and Battjes 1985). Within a year, two additional reviews were published describing the application of this approach to alcohol and illicit drugs (Battjes 1985; Botvin 1986). Of the 20 or so studies discussed in these papers, just two—Botvin et al. (1984) and McAlister et al. (1979)—presented data pertaining to the effects of social influence programs on illicit drugs (in each case marijuana).²

In addition to the studies cited in these reviews, there were four other accounts of the effects of social influence programs on illicit drug use available at the time—three from the United States (DuPont and Jason 1984;

Moskowitz, Malvin, et al. 1984; Moskowitz, Schaps, et al. 1984) and one from New Zealand (Casswell, Mortimer, and Gilroy 1982). The results of all six studies are summarized in Table 1.

The findings of the two favorable studies are far from compelling. McAlister et al. (1979) found that 7.6% of seventh grade students who received an RST program reported smoking marijuana in the past week or day, compared to 14.9% of those in a comparison school ($p < .01$). However, because data on marijuana use were only reported at follow-up, it is impossible to rule out that students from the two schools were different prior to the intervention. In the evaluation described by Botvin et al. (1984), seventh grade students received either a 20-session SST program delivered by classroom teachers or the same program delivered by "peer leaders." Compared to students in a nonintervention comparison condition, there were significantly fewer students using marijuana in the peer-led SST group at posttest, but no differences between the teacher-led group and the comparisons. At a subsequent 1-year follow-up, there was again no statistically significant differences between the teacher-led SST condition and the comparison condition. Moreover, the effects of the peer-led program were patchy. Five outcome variables were assessed (ever used, monthly use, weekly use, use in previous 24 hours, and a 5-point index combining all of the scales). In addition, during the intervening year, subjects in one of the peer-led groups received a 10-session booster. Of the 10 comparisons made between the peer-led groups and the comparison group at 1 year (2 study conditions \times 5 outcome measures), only two were statistically significant (peer-led booster group monthly recall and index measure).

The other four studies shown in Table 1 found no statistically significant differences in patterns or levels of illicit drug use between recipients of social influence programs and comparison subjects at follow-up. The bulk of available evidence therefore indicated that social influence programs were little better than earlier programs. In short, by 1986, when the federal government committed more than \$200 million to school-based programs to fight illicit drug use, evidence indicating the effectiveness of this strategy was almost nonexistent.

THE RESULTS

During the early days of the War on Drugs, data on adolescent drug use from large national surveys were used by policy makers to argue that drug prevention strategies were working. For example, the U.S. Senate Committee

TABLE 1: Summary of Evaluations of School-Based Social Influence Programs, Pre-1986

Authors	Conditions ^a	Number of Sessions ^b	Baseline n ^c (% followed up)	Follow-Up in Months	Type of Comparison ^d
McAlister et al. (1979)	RST Comp	8	Not reported	21	NR
Summary of results. Significantly fewer of the students in the intervention school reported use of marijuana at follow-up than students in the comparison school. However, there was no baseline assessment, so it is impossible to know whether differences in marijuana use levels existed between the two schools prior to the intervention.					
Botvin et al. (1984)	SST (teacher) SST (peer) Comp	20 + 10	1,311 (90%)	PT	RA-school
Summary of results. At posttest, significantly fewer students in the peer-led intervention group than in the comparison group recalled using marijuana during the previous month and previous week. There were no statistically significant differences between the teacher-led intervention group and the comparison group.					
Casswell, Mortimer and Gilroy (1982)	RST Comp	6	1,931	3	RA-school
Summary of results. "There were no demonstrable effects of the education programme on self-reported drug use and anticipated future use among this sample of high school students" (Casswell, Mortimer, and Gilroy 1982, 351).					
DuPont and Jason (1984)	RST Information Comp	9	Not reported	2	RA-student
Summary of results. Results for use of marijuana/hashish not reported separately from alcohol and cigarettes due to low frequency of reported use. Analysis of combined drug use showed "no significant changes in drug use across testing points" for the RST group (DuPont and Jason 1984, 374).					

<p>Moskowitz et al. (1983), Moskowitz, Schaps, et al. (1984)</p>	<p>SST^e Comp</p>	<p>12 687 (69%)</p>	<p>PT, 12</p>	<p>NR-school</p>
<p>Summary of results. Twenty-six comparisons were made between the SST and comparison groups for both male and female subjects. Six significant differences for female subjects at both posttest and 12-month follow-up. None pertained to use of illicit drugs. Three significant differences for male subjects at posttest and 12-month follow-up. Again, none pertained to illicit drug use.</p>				
<p>Moskowitz, Malvin, et al. (1984)</p>	<p>SST Comp</p>	<p>12 473 (74%)</p>	<p>PT, 5</p>	<p>RA-school</p>
<p>Summary of results. "The course was found to have had no significant effect on girls and only a few effects at follow-up for boys" (Moskowitz, Malvin, et al. 1984, 9). None of these effects pertained to the questionnaire item "involvement in marijuana use."</p>				
<p>a. RST = resistance skills training; SST = social skills training; Comp = comparison group. Titles of other programs are listed in full. b. Most sessions were 45 to 60 minutes long. Where N₁ + N₂ appears, the later number refers to booster sessions delivered after the main program. c. In two papers, the pretest sample size is not reported; rather, the authors state the number who supplied data at all measurement points. The size of these "analysis samples" are 526 in the study of McAlister et al. (1980) and 58 in the study of DuPont and Jason (1984). d. All studies include a comparison group. RA refers to random allocation and NR to nonrandom allocation. Details are also presented of the unit used in allocating subjects to conditions, for example, by student, by classroom, or by school. e. This course included a teacher training component and an alternatives program.</p>				

on Labor and Human Resources (1990) stated that the decline in reported use of cocaine and marijuana, along with the increase in perceptions of harm associated with these drugs, evident in the *Monitoring the Future Study* was a sign of the effectiveness of drug prevention efforts. Two years later, the introduction to the annual *National Drug Control Strategy* returned to the idea of a two-front war and declared that "the first front is against casual use, and we are winning. For those who are younger, and especially adolescents, there is only good news. Drug use is down substantially for these groups during the last several years, showing that our efforts are, in effect, shutting down the pipeline and preventing the entry of new users" (Office of National Drug Control Policy 1992, 4).

The good news ceased in 1993, however, as national surveys showed an increase in adolescent drug use for the first time in more than a decade. This upward trend continued during subsequent years. In the *National Household Survey*, monthly marijuana use among 12- to 17-year-olds increased from 4% in 1992 to more than 7% in 1994, whereas perceived risks of use declined. *Monitoring the Future* showed that this trend was evident among 8th, 10th, and 12th graders. Among the latter, reported use of any illicit drug during the previous 30 days rose from 14.4% in 1992 to 23.8% in 1995. During the same time period, the proportion who disapproved of occasional use of marijuana fell from 80% to 67%, and the proportion who thought that occasional use was potentially harmful from 40% to 26% (Johnson, O'Malley, and Bachman 1996).

THE RESPONSE

These survey results have presented advocates of current policies with a clear dilemma—how could the present approach be a success in the face of the most basic evidence indicating otherwise? The response by advocates of school-based prevention programs has focused on two issues: First, it is argued that drug use and accompanying favorable attitudes have increased among young people as the financial commitment by federal government to drug prevention has declined; and second, it is argued that these problems have become worse because the right types of prevention programs are not being supported.

MORE MONEY, LESS DRUG USE

Advocates of drug prevention argue that there is a causal relation between cuts in federal spending and increased drug use among youth. For example,

in 1996, Education Secretary Richard Riley, responding to proposed reductions in the drug prevention budget argued:

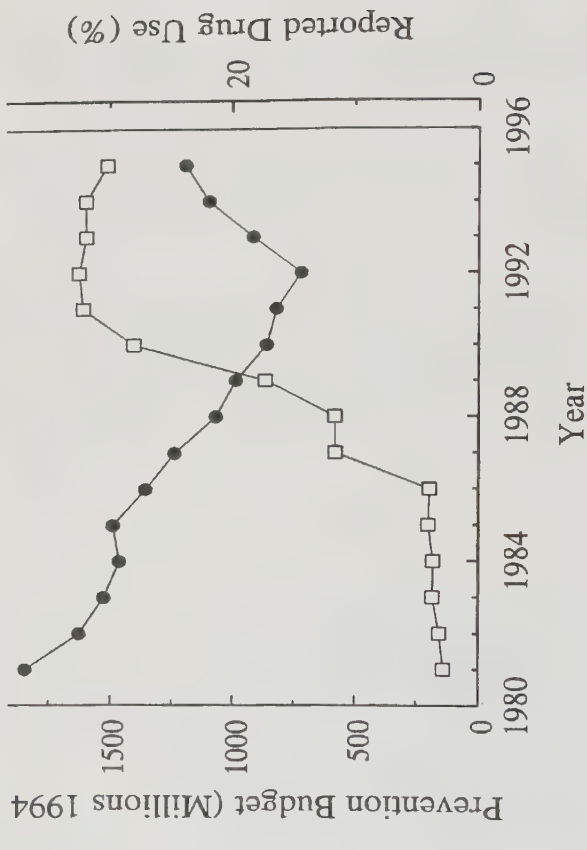
This retreat from federal support comes at a time when drug use among young people is rising. . . . It's a very disturbing and clear signal that we must redouble our efforts. I cannot understand why anyone concerned with the future of the nation and the health and safety of our children would willingly retreat in such a critical fight. ("Riley Details" 1996, P. 4)

Researchers, too, have linked the increase in adolescent drug use to recent stagnation in federal funding for prevention programs. Lloyd Johnston (Wren 1996), principal investigator on *Monitoring the Future*, recently observed:

Each new generation needs to learn the same lessons about drugs if they're going to be protected from them. . . . Unless we do an effective job of educating the newer generations, they're going to be more susceptible to using drugs and have their own epidemic. And I think that's what is happening now. (P. A11)

Elsewhere, I have discussed at length the merits of the argument that increased adolescent drug use is in any way the result of reduced federal funding for drug prevention activities (Gorman 1997) and therefore will not go into detail here. Suffice it to say that, as Figure 1 makes clear, drug use was declining for a number of years before the massive federal buildup in prevention spending. The figure presents data on the annual federal budget for drug use prevention and prevalence of illicit drug use among 12th graders for the period 1981 (the first year for which Office of National Drug Control Policy data on spending are available) to 1995 (the year for which most recent data are available). Advocates of current policies and programs like to begin their history of drug prevention with 1986 as the base year. From this perspective, 5 years of increased spending coincides with 5 years of declining drug use (1987-1995), and 3 years of reduced spending coincides with 3 years of increased drug use (1992-1995). This version of history ignores the fact that the decline in adolescent drug use began in 1979 and that the rate of decline was virtually the same during the years of modest federal spending on drug prevention as during the years of high federal spending.

It is, of course, difficult to demonstrate either the success or failure of drug prevention policies from the type of aggregate-level drug use data shown in Figure 1 (although, as noted above, politicians and others frequently try to do so). Numerous factors affect drug use behavior, and such behavior may be an insensitive indicator of the effectiveness of our prevention efforts. However, one indicator that might be more sensitive to the effects of drug



● = Drug Use (Source: Johnson, O'Malley and Bachman 1996, Table 13).
 □ = Prevention Budget (Source: Office of National Drug Control Policy 1996, pp. 316-319; converted to millions of constant 1994 dollars).

Figure 1: Federal Government Spending on Drug Prevention and Reported Drug Use by 12th Graders, 1981-1995

education efforts in schools is young people's attitudes toward drug use. Whatever their potential effects on behavior, even the most insipid school-based prevention program is explicitly intended to persuade young people that drug use is wrong. Clearly, other areas of drug control policy (e.g., treatment or interdiction) are not intended to have such a direct effect upon adolescent attitudes.

Table 2 shows the relationship between federal spending on school-based drug prevention and 12th graders' attitudes toward occasional marijuana use between 1981 and 1995. As with drug use behavior, the trend for attitudes was in the right direction (i.e., downward) long before the federal government decided to create drug-free schools in 1986, and the proportion expressing disapproval of occasional drug use continued to increase during the next 5 years of accelerated federal spending on school-based programs (1987-1991). However, the rate of increase was no better after the infusion of hundreds of millions of dollars than before: From 1981 to 1985, the proportion who disapproved of occasional use rose by 13.8%, compared to 7.8% for the period 1987-1991. Thus, at this level of aggregation, increased spending on school-based drug education does not appear to have met one of its most basic goals—to persuade young people that drug use is wrong.

BETTER DISSEMINATION OF "BETTER" PROGRAMS

Among prevention researchers, the response to the recent increase in adolescent drug use has been to call for more effective diffusion of those programs purportedly shown to be most effective in reducing drug use (Dusenbury and Falco 1995). And, as in the mid-1980s, it is the social influence approach that is said to hold the most promise. Rohrbach et al. (1996), for example, state that despite "promising evidence of effectiveness," these programs have yet to be widely adopted by schools. They add: "Because young people are not being exposed to the psychosocial-based programs that research has shown to be effective, the public health impact of these strategies has been minimal. More effective diffusion of these programs is essential if their impact is to be increased" (p. 921).

As Brown and Horowitz (1993) observe, perceptions concerning the effectiveness of social influence programs have been shaped to a considerable degree by the findings of a handful of large-scale evaluation studies. Notable among these are evaluations of the Life Skills Training (LST) program developed by Gilbert Botvin, Project SMART developed by William Hansen, and Project ALERT developed by Phyllis Ellickson. Table 3 summarizes

TABLE 2: Disapproval of Occasional Marijuana Use Among 12th Grade Students and Department of Education Drug Prevention Budget, 1981-1995

	Year															
	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	
% who disapprove of occasional marijuana use ^a	52.6	59.1	60.7	63.5	65.8	69.0	71.6	74.0	77.2	80.5	79.4	79.7	75.5	68.9	66.7	
Department of Education Budget (in \$ millions) ^b	4.7	4.5	4.3	4.1	4.1	3.9	263.9	287.3	425.4	612.1	663.9	663.7	619.4	490.6	473.2	
% increase/decrease over previous year	—	-4	-4	-5	0	-5	+6,667	+9	+48	+44	+9	0	-7	-21	-4	

a. Johnson, O'Malley, and Bachman (1996, Table 22); percentage reporting that they "disapprove of people (who are 18 or older) [who] smoke marijuana occasionally."

b. Office of National Drug Control Policy (1996, 316-19); converted to millions of constant 1994 dollars.

findings pertaining to illicit drug use from the major evaluations of these programs.³ Below, the projects are discussed in more detail.

Project SMART

One of the first large-scale evaluations of the effects of a social influence program on illicit drug use was conducted by Hansen and colleagues in eight junior high schools in Los Angeles and involved almost 3,000 seventh grade students (Hansen et al. 1988). A 12-session RST program was delivered by health educators and regular teachers during the course of one semester. Separate analyses were conducted for those who were present at baseline and 12-month follow-up, and for those present at baseline and 24-month follow-up. Attrition was high—37% at 12 months and 52% at 24 months. In assessing the effects of Project SMART, two types of analyses were performed: one involving only those subjects who reported no marijuana use in the 30 days prior to baseline (“non-users”), and one involving all subjects irrespective of baseline use. Data were presented in terms of, first, the proportion of subjects who changed their level of use at follow-up, and, second, scores on an index measuring average number of marijuana joints per student per week.

In assessing onset among baseline nonusers, outcome was presented in the form of different levels of use ranging from “one time or more” through to “21 or more times.” The only statistically significant difference (at the conventional level of $p < 0.05$) between SMART and comparison group subjects at either 12- or 24-month follow-up occurred at the level of “one time or more” (7% versus 11%). No significant differences were found when higher levels of use were considered (i.e., anything greater than one time), and even the effect on this low level use was not evident at 24 months. On the index measuring joints smoked per week, there were no statistically significant differences between the intervention and control groups at either follow-up.

When all subjects, irrespective of baseline use, were considered, there were no significant differences (again, at the conventional 0.05 level) between the SMART and comparison group on the index of joints smoked per week. There were also no statistically significant differences between the two groups in terms of the proportion who either increased or decreased use at each follow-up.

As Hansen (1995) observes, the SMART curriculum became a guide to the “best” school-based program components and was the prototype for a number of other curricula, including D.A.R.E. The results summarized above show that its effects on marijuana use were minimal at best: It simply delayed

TABLE 3: Summary of Major Evaluations of School-Based Social Influence Programs, Post-1986

<i>Authors</i>	<i>Conditions^a</i>	<i>Number of Sessions^b</i>	<i>Baseline n (% followed up)</i>	<i>Follow-Up in Months</i>	<i>Type of Comparison^c</i>
Hansen et al. (1988)	RST				
	Affective Comp	12	2,863 (48%)	12, 24	RA-school
Summary of results. "The results of this study to date suggest that the most efficacious program for reducing drug onset is a program that features social influences resistance training as a major focus" (Hansen et al. 1988, 150).					
Hansen and Graham (1991)	RST				
	Normative				
	RST + Normative Information	9	3,011 (71%)	12	RA-classroom
Summary of results. "[The] program designed to teach seventh-grade students techniques to refuse offers to use substances had no discernible positive impact on substance use behavior" (Hansen and Graham 1991, 425).					
Ellickson and Bell (1990); Ellickson, Bell, and McGuigan (1993)	RST-peer RST-teacher Comp	8 + 3	6,527 (54%)	3, 12, 15, 36, 60	RA-school
Summary of results. On short-term effects: "Project ALERT's results during the junior high school years provided strong evidence that the social influence approach to prevention can curb . . . marijuana use" (Ellickson, Bell, and McGuigan 1993, 858). On long-term effects: "By the end of high school, the treatment program no longer had a significant effect on behavior" (Ellickson, Bell, and McGuigan 1993, 859).					
Botvin et al. (1994); Botvin, Schinke, et al. (1995)	SST Culturally focused Information	15 + 8	757 (60%)	PT, 24	RA-school
Summary of results. Prevalence of marijuana use too low at posttest (1%) to conduct analyses. At the 2-year follow-up, the "proportion of marijuana experimenters was similar across conditions" (Botvin, Schinke, et al. 1995, 189). The proportions for each condition were 19% (SST), 18% (culturally focused), and 20% (information).					

Botvin, Baker, Dusenbury, et al. (1990); Botvin, Baker, et al. (1995)	SST (personal training) SST (video training) Comp	20 + 10	5,954 (60%)	36, 72	RA-school
-----------------------------------------------------------------------------	---------------------------------------------------------	---------	-------------	--------	-----------

Summary of results. "The results of this 6-year randomized trial provide additional evidence that school-based drug abuse prevention programs in general and this approach in particular can reduce the prevalence of drug use" (Botvin, Baker, et al. 1995, 1111).

- a. RST = resistance skills training; SST = social skills training. Comp = comparison group. Titles of other programs are listed in full.
b. Most sessions were 45 to 60 minutes long. Where $N_1 + N_2$ appears, the later number refers to booster sessions delivered after the main program.
c. All studies include a comparison group. RA refers to random allocation and NR to nonrandom allocation. Details are also presented of the unit used in allocating subjects to conditions, for example, by student, by classroom, or by school.

low level use for 1 year among baseline nonusers. There were no long-term effects among this group, and no effects at all among those who had already initiated use at baseline. Moreover, as Table 2 shows, a subsequent evaluation of a similar RST program by Hansen and Graham (1991) found no effects on marijuana use.

Project ALERT

Project ALERT formed the basis of the national drug prevention initiative known as the BEST Campaign, and was cited by the U.S. Senate Committee on Labor and Human Resources (1990) as "a documented success in the effort to reduce the demand for drugs through education" (p. 33). In the large-scale evaluation of the program carried out in Oregon and California, seventh grade students from 30 junior high schools were randomly allocated to one of three conditions—a health educator-led RST program, a teacher-led/peer-assisted RST program, and a nonintervention comparison group. The program comprised eight sessions during the first year, with three booster sessions during the second year. In the initial phase of the evaluation, follow-ups were conducted at 3, 12, and 15 months after the intervention (Ellickson and Bell 1990). Long-term effects were assessed through further follow-ups conducted in Grades 10 and 12 (Ellickson, Bell, and McGuigan 1993). Again, attrition was a problem in the study: Of more than 6,500 individuals who were assessed at baseline, fewer than 4,000 were included in subsequent analyses.

In the data analyses, the sample was broken down into three risk groups according to baseline drug use. In the case of marijuana, these groups were based on prior use of the drug and prior use of cigarettes—nonusers of both (low risk), marijuana nonusers/cigarette users (moderate risk), and users of both (high risk). At the 3-, 12-, and 15-month follow-ups, the effects of the program were assessed for five specific outcome variables—ranging from "ever used" to "weekly use," and including "quitting" among baseline users. This combination of experimental conditions, risk groups, follow-up periods, and outcome variables resulted in 68 logically possible comparisons between those receiving ALERT and those not. Of these, just six were statistically significant at the traditional level of $p < 0.05$ (Gorman 1994).

Ellickson and Bell (1994) have argued that it is unfair to judge ALERT in these terms, as some of the comparisons involve subgroups displaying too little drug use for meaningful statistical analysis. However, even by their own standards, the effects on marijuana use were unimpressive: Of the 38 comparisons they made in their 1990 article between ALERT subjects and comparisons in the high and moderate risk groups, only two were statistically

significant. The remaining four differences occurred among the low-risk group and were limited mainly to the outcome "ever used" rather than measures of more intense use (e.g., monthly use).

Differences between ALERT subjects and comparisons were nonexistent at the 10th and 12th grade follow-up. Ellickson, Bell, and McGuigan (1993) attribute this to the absence of booster sessions in the schools after the first year of the program, and they call for additional research to develop and test such efforts. This ignores the fact that the short-term effects of ALERT were minimal. Why would high and moderate risk subjects benefit from more of the program? Why is more evaluation required? As Ellickson (1995) observes, booster sessions are intended to "extend program effects." For ALERT, there were essentially no program effects to extend.

The latter aspect of the ALERT evaluation illustrates a peculiar feature of school-based drug prevention research during the past 10 years: Whatever the outcome, the recommendation is for more of the program and more evaluation. With the exception of D.A.R.E., negative findings are seldom accompanied by a suggestion that we try something else. Information and affective programs of earlier years were unable to survive negative evaluations; in contrast, social influence programs invariably live to fight another day.

Life Skills Training Program

A further example of this tendency of social influence programs to thrive in the face of weak research findings is provided by the development of research into the effects of the Life Skills Training (LST) program in urban settings. An early pilot study of the program's effect on smoking among African American youths showed virtually no influence on behavior (statistically significant results were present for only one of seven variables measured) and little effect on hypothesized mediating variables such as attitudes and social skills (Botvin, Batson, et al. 1989). A second pilot study, again concerned with smoking prevention but targeted this time at Hispanic students, also produced very patchy results (Botvin, Dusenbury, et al. 1989). Program effects on smoking behavior approached statistical significance only in the case of smoking during the past month, but not for smoking during the past week or past day or intentions to smoke in the future. Effects on hypothesized mediating variables were found for scales assessing knowledge and attitudes, but not social skills or psychological factors. The reports of both pilot studies concluded that the results provided evidence for the efficacy of the LST program with urban youth and suggested that a large-scale study would serve to demonstrate statistically significant effects.

A subsequent large scale evaluation of the effects of the program on smoking among more than 3,000 seventh grade students from 47 schools in New York City found statistically significant differences in only one of five behavioral measures (Botvin et al. 1992). Significant differences between groups at follow up were also reported on measures of knowledge and normative expectations, but the magnitude of these were very small and of questionable practical significance (see Gorman 1995). Despite these results, the study was said to have extended the results of previous research and to have demonstrated the "generalizability of this approach to predominantly Hispanic urban minority students" (Botvin et al. 1992, 290).

Results of the only published evaluation of the LST program to assess its effects on marijuana use among minority students (Botvin, Schinke, et al. 1995) are shown in Table 2. In this study of 456 seventh grade students from six public schools in New York City, the proportion reporting experimenting with the drug at 2 year follow up were virtually identical across study conditions, and scores on an index of marijuana use frequency were also the same.

It is curious that, despite the fact that there is no evidence showing that the LST program prevents use of illicit drugs among urban minority youth, and that its effects on cigarette smoking are limited at best to low level experimental use, the program is recommended with enthusiasm to grantees by the federal agencies concerned with developing interventions for this target population (Center for Substance Abuse Prevention 1993a, 1993b; National Institute on Drug Abuse, 1997). For their part, the developers of the LST program consider the evidence from school- and community-based evaluations of social influence programs to be sufficiently compelling to state "It is now incumbent upon health care professionals, educators, community leaders, and policymakers to move expeditiously toward [their] wide dissemination and utilization" (Botvin and Botvin 1992, 924). Such claims can best be judged, in the case of the LST program, through consideration of the largest evaluation published to date (Botvin, Baker, Dusenbury, et al. 1990; Botvin, Baker, et al. 1995). This long term follow-up study has been hailed by the popular and professional press alike as providing convincing evidence that effective school-based drug prevention exists (Dusenbury and Falco 1995; Mathias 1994; Van Biema 1996).

The 6-year follow up study of the LST program was conducted in 80 schools in the state of New York, with predominantly White, middle-class seventh grade students. Two experimental conditions—one in which training in the use of LST was conducted through a 1-day workshop and one in which training was provided through a videotape and written material—were compared with a no intervention comparison condition. Students in the intervention conditions received 15 LST classes in seventh grade, 10 in eighth grade,

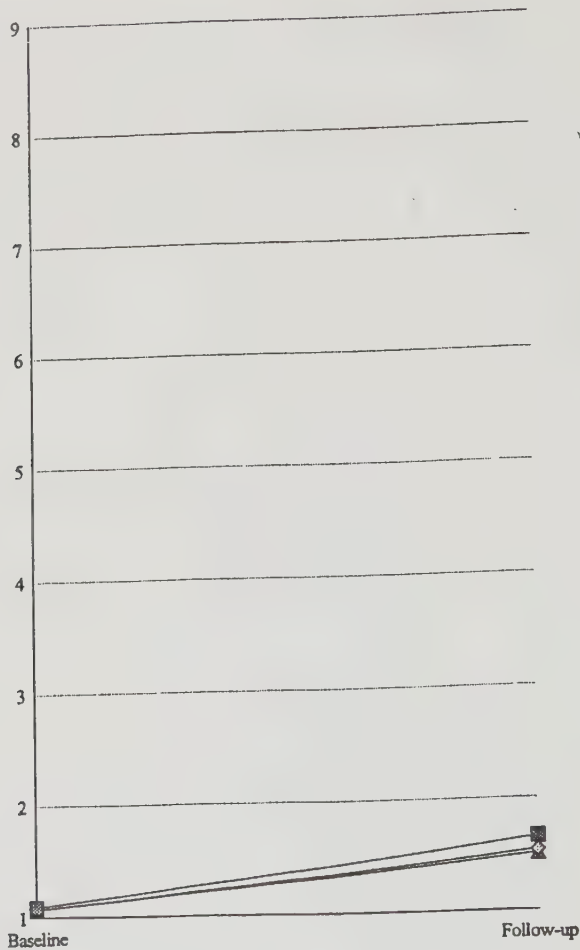
and 1 in each group. The effects of the program were assessed at 24 months (Green et al. 1974) and 36 months (Green, Koser et al. 1975).

In the first study, nine children were included in a group of 1504 children. 43% of these originally included in the study later (1974) and 41% of those for whom data were available at both baseline and follow-up (n = 1,066). To be included in the "high-fluoride sample," children had to have 100% on a measure of program implementation. The criterion data were reported for those included from the high-fluoride sample. Data were collected through self-completed questionnaires, and treatment was not assigned on a 2-point frequency scale ranging from never (0) to every day (2). A few of the following statistically significant differences were found between the water intake of the two treatment groups and the comparison group. The mean of the first group is given in Figure 2 in an effort to understand the general significance of these findings.

In the figure shown, the mean of first group was calculated as mean of two subgroups. The two subgroups groups had a mean value of 1.07 on the intake and the comparison group 1.05. By the time of the 2-year follow-up, mean of each group of the first was 1.15 and 1.04 compared to 1.16 for the second group. It was at 2 on the intake information that the difference of 1.07 was found. Thus, the mean value of all groups increased around the last end of the study indicating little use of the program. The difference between the 1.17 group and the comparison group was mean value of 0.7 on a 2-point scale.

In the 2-year evaluation, a really different situation was found in terms of the program coverage (Green et al. 1975). The intervention was not identified as lower of two subgroups because frequency measure ("never" and "every day") was taken program effect data provided for all children included at baseline and follow-up in the "high-fluoride" (1.17) sample as 100% of the baseline sample. No statistically significant differences between the 1.17 group and the comparison were found. Following this situation, were included in a high-fluoride sample program of children who had received a sufficient amount of the program as defined in the 1975 study, and program effect data were not collected. The average use of monthly program effect data was 100% for the 1.17 group and 90% for the comparison group (p < 0.05, one-tailed test). The average use for monthly use was 94% for the 1.17 group and 89% for the comparison group (p < 0.05, one-tailed test).

The program with these studies is that there are based on a comparison of a selected subgroup of treatment group subjects with a comparison group that intervened or not intervened. Figure 3 presents details of the manner in which the high-fluoride sample was formed. Of the 1,450 1.17



- ▲ = Life Skills Training program with teachers taught through workshop.
- ◆ = Life Skills Training program with teachers taught through videotape.
- = Comparison Group.

Source: Botvin et al. (1990b), Table 2 and Table 3.

Figure 2: Mean Baseline and Follow-Up Scores on an Index of Marijuana Use of Study Groups in Botvin, Baker, Dusenbury, et al.'s (1990) 3-Year Follow-Up Evaluation of the Life Skills Training Program

subjects included in the full sample at the 6-year follow-up, 845 (34%) were excluded from the high-fidelity subsample. Indeed, only about 4 of every 10 LST subjects assessed at baseline were eventually included in the high-fidelity sample 6 years later.⁴ Although the high-fidelity and full samples were virtually identical in terms of demographic characteristics such as gender and race, it simply cannot be ruled out that the two groups differ in some fundamental way that affected the dosage of program they received. This difference could exist at the level of the individual subject (e.g., motivation, level of school attendance) and/or at the level of the classroom or school (e.g., interest of teachers in drug prevention, ability to deliver the program competently). In each case, such factors could affect not only program dosage but also reported drug use—for example, subjects motivated not to use drugs are likely to be more conscientious about program attendance. In short, the differences found using the high-fidelity subsample might simply be spurious, due not to program content, but to self-selection of subjects and/or their schools or classrooms into the intervention. Reviews and reports of the study, such as those by the National Institute on Drug Abuse (1997), ignore the fact that the positive findings concerning illicit drug use are limited to the high-fidelity sample and never raise the fundamental issue of selection into treatment conditions.

As the above overview of LST studies indicates, evaluations will nearly always reveal some differences between those who receive an intervention program and those who do not. Thus, as Botvin observes, “Depending on the measure used,” the evaluations can be interpreted as “providing further support for the effectiveness of the LST prevention approach” (Botvin 1996, 229). However, if other outcome measures included in the studies “are used” to assess effectiveness, they provide little support for the continued use of this approach.

CONCLUSION

The evidence presented herein, from both national surveys and program evaluations, shows that we have yet to develop successful techniques of school-based drug prevention. The claims made on behalf of this aspect of the nation’s drug control policy are largely unsupported by empirical data. Evidence is cited selectively to support the use of certain programs, and there is virtually no systematic testing of interventions developed in line with competing theoretical models of adolescent drug use. As I have observed elsewhere, theory testing in the field of drug prevention has been conducted

	<u>LST-E1^a</u>	<u>LST-E2^b</u>	<u>Comparison</u>	<u>Total</u>
6-year high fidelity sample	762	848	1142	2752
	↑ (68%)	↑ (64%)	↑ (100%)	↑ (77%)
6-year full sample	1128	1327	1142	3597
	↑	↑	↑	↑ (81%)
3-year full sample	3037 ^c		1429 ^c	4466
	↑		↑	↑ (75%)
Baseline sample	4049 ^c		1905 ^c	5954

Figure 3: Details of Samples Used in Botvin, Baker, Dusenbury, et al. (1990) and Botvin, Baker, et al. (1995) Long-Term Follow-Up Study of the Life Skills Training Program

SOURCE: Botvin, Baker, Dusenbury, et al. (1990, 349) and Botvin, Baker, et al. (1995, 1107, 1110).

a. LST-E1 = Life Skills Training program with teachers taught through workshop.

b. LST-E2 = Life Skills Training program with teachers taught through videotape.

c. Estimated (see Note 4 for details).

using an inductive methodology, in which the function of research is to accumulate "confirming instances" of program effectiveness (Gorman 1996). This task is easily achieved as evaluations can be structured so as to ensure positive results by, for example, measuring numerous outcome variables. Alternatively, in the face of nonsupportive evidence, data sets can be modified (e.g., by focusing on specific subsamples of subjects) or the criteria for success altered (e.g., from behavior change to change in attitudes or knowledge). Policy makers have, for the most part, uncritically accepted—indeed encouraged—such research.

The question remains as to why policy makers champion drug prevention programs that have so little grounding in empirical research. In considering this, it is instructive to recall that for close to 30 years, Soviet agricultural policy was developed in accordance with the theories and research of Trofim Lysenko. According to Lysenko's theory of inherited acquired characteristics, it was possible to transform one crop into another (e.g., wheat into rye) through changing its environment (e.g., by planting it in a different season). Lysenko's "science" thrived under Stalin's regime, in the face of disastrous consequences, as it was totally in accord with the prevailing political philosophy; research data were irrelevant. Similarly, the belief that school-based programs can teach children the skills to be "drug free" is entirely in keeping with the individually orientated, zero-tolerance orthodoxy of current U.S. drug control policy. The programs thrive not because research demonstrates their efficacy and superiority over competing approaches, but because the principles upon which they are based are compatible with the prevailing wisdom that exists among policy makers and politicians. And, judging from recent government publications and the viciousness with which critics are attacked, the uncritical acceptance of school-based social skills training seems likely to continue into the near future.

There are, however, a few positive signs, such as the recent call by a group of prominent drug policy analysts and researchers for an infusion of reason and better judgment into the discussion of current policies (Wren 1997). It is hoped that the latter effort can move beyond the usual focus on interdiction and law enforcement to a reevaluation of all aspects of U.S. drug control policy, including prevention. This would entail assessing the full range of evidence concerning social influence programs, including studies that indicate potentially harmful program effects especially among those most at risk (Brown and D'Emidio-Caston 1995; Palinkas et al. 1996). To do otherwise, and continue to advocate the use of school-based social influence programs on the basis of selected, isolated positive findings, is in the interest of no more than a very few individuals. For, as Daniel Patrick Moynihan observed of a key component of another of America's social policy wars, government

intervention in social problems, while necessary, is also risky and uncertain. "It requires," he added, "enthusiasm, but also intellect, and above all it needs an appreciation of how difficult it is to change things and people. Persons responsible for such programs who do not insist on clarity and candor in the definition of objectives, and the means for obtaining them . . . do not much serve the public interest" (Moynihan 1966, 8).

NOTES

1. All federal drug control budgets are from the Office of National Drug Control Policy (1996) and have been converted to constant 1994 dollars.

2. Botvin (1986) also cites two unpublished reports—McAlister (1983) and Botvin et al. (1985). The latter is, in all likelihood, a 1-year follow-up, as it has virtually the same title as a published account of the effects of the intervention at 1 year (Botvin, Baker, Filazzola, et al. 1990; see text for discussion).

3. The table excludes the D.A.R.E. program, which is similar in content to the programs shown. D.A.R.E. evaluations are conducted by individuals independent of the process of program design, implementation, and marketing, and show that the program has no effect on drug use (Rosenbaum et al. 1994; Clayton, Cattarello, and Johnstone 1996). The table also excludes the Mid-Western Prevention Project (Pentz et al. 1989), a school-based program with additional elements such as mass media. Its limitations are discussed in detail elsewhere (Aguirre-Molina and Gorman 1996; Brown and Horowitz 1993; Gerstein and Green 1993; Gorman and Speer 1996).

4. Neither published account gives details of the number of subjects in the LST and comparison groups at baseline. The numbers shown in Figure 3 (4,049 and 1,905, respectively) are based on the assumption that attrition at the 3-year follow-up was similar in both groups (i.e., 25% in each). An alternative way to derive the number of subjects in each condition is to base the calculation on the number of schools in each. At outset, 56 schools were recruited, 34 (61%) of which were assigned to the LST group and 22 (39%) to the comparison group. Using these proportions to estimate the number of subjects in study conditions, there are about 3,600 LST students and about 2,300 comparisons at outset. In this case, the high-fidelity sample at the 6-year follow-up represents about 45% of the original sample (1,610/3,600).

REFERENCES

- Aguirre-Molina, M. and D. M. Gorman. 1996. Community-based approaches for the prevention of alcohol, tobacco and other drug use. *Annual Review of Public Health* 17:337-58.
- Anti-Drug Abuse Act. 1986. Public Law 99-570. Washington, DC: U.S. Government Printing Office.
- . 1988. Public Law 100-690. Washington, DC: U.S. Government Printing Office.
- Battjes, R. J. 1985. Prevention of adolescent drug abuse. *International Journal of the Addictions* 20:1113-34.

- Bell, C. S. and R. Battjes. 1985. *Prevention research: Deterring drug abuse among children and adolescents*. NIDA Research Monograph 63. Rockville, MD: NIDA.
- Bennett, W. J. 1989. Introduction. In *National Drug Control Strategy, 1989*. Washington, DC: The White House.
- . 1992. *The de-valuing of America: The fight for our culture and our children*. New York: Simon & Schuster.
- Berberian, R. M., G. Gross, J. Lovejoy, and S. Paparella. 1976. The effectiveness of drug education programs: A critical review. *Health Education Monographs* 4:377-98.
- Botvin, G. J. 1986. Substance abuse prevention research: Recent developments and future directions. *Journal of School Health* 56:369-74.
- . 1996. Substance abuse prevention through life skills training. In *Preventing childhood disorders, substance abuse, and delinquency*, edited by R. DeV. Peters and R. J. McMahon, 215-40. Thousand Oaks, CA: Sage.
- Botvin, G. J., E. Baker, L. Dusenbury, E. M. Botvin, and T. Diaz. 1995. Long-term follow-up results of a randomized drug abuse prevention trial in a White middle-class population. *Journal of the American Medical Association* 273:1106-12.
- Botvin, G. J., E. Baker, L. Dusenbury, S. Tortu, and E. Botvin. 1990. Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: Results of a 3-year study. *Journal of Consulting and Clinical Psychology* 58:437-46.
- Botvin, G. J., E. Baker, A. Filazzola, E. Botvin, et al. 1985. *A cognitive-behavioral approach into substance abuse prevention: A one-year follow-up*. Paper read before the 93rd annual meeting of the American Psychological Association, 23-27 August, Los Angeles.
- Botvin, G. J., E. Baker, A. D. Filazzola, and E. M. Botvin. 1990. A cognitive-behavioral approach to substance abuse prevention: One-year follow-up. *Addictive Behaviors* 15:47-63.
- Botvin, G. J., E. Baker, N. L. Renick, A. D. Filazzola, and E. M. Botvin. 1984. A cognitive-behavioral approach to substance abuse prevention. *Addictive Behaviors* 9:137-47.
- Botvin, G. J., H. W. Batson, S. Witts-Vitale, V. Bess, E. Baker, and L. A. Dusenbury. 1989. A psychosocial approach to smoking prevention for urban Black youth. *Public Health Reports* 104:573-82.
- Botvin, G. J. and E. M. Botvin. 1992. School-based and community-based prevention approaches. In *Substance abuse: A comprehensive textbook*, edited by J. H. Lowinson, P. Ruiz, and R. B. Millman, 910-27. Philadelphia: Williams and Wilkins.
- Botvin G. J., L. Dusenbury, E. Baker, S. James-Ortiz, E. Botvin, and J. Kerner. 1992. Smoking prevention among urban minority youth: Assessing effects on outcome and mediating variables. *Health Psychology* 11:290-99.
- Botvin, G. J., L. Dusenbury, E. Baker, S. James-Ortiz, and J. Kerner. 1989. A skills training approach to smoking prevention among Hispanic youth. *Journal of Behavioral Medicine* 12:279-96.
- Botvin, G. J., S. P. Schinke, J. A. Epstein, and T. Diaz. 1994. Effectiveness of culturally focused and generic skills training approaches to alcohol and drug abuse prevention among minority youths: Two-year follow-up results. *Psychology of Addictive Behaviors* 8:116-27.
- Botvin, G. J., S. P. Schinke, J. A. Epstein, T. Diaz, and E. M. Botvin. 1995. Effectiveness of culturally focused and generic skills training approaches to alcohol and drug abuse prevention among minority adolescents: Two-year follow-up results. *Psychology of Addictive Behaviors* 9:183-94.
- Braucht, G. N., D. Follingstad, D. Brakarsh, and K. L. Berry. 1973. Drug education: A review of goals, approaches and effectiveness, and a paradigm for evaluation. *Quarterly Journal of Studies on Alcohol* 34:1279-92.

- Brown, J. H. and M. D'Emidio-Caston. 1995. On becoming at-risk through drug education: How symbolic policies and their practices affect students. *Evaluation Review* 19:451-92.
- Brown, J. H. and J. E. Horowitz. 1993. Deviance and deviants: Why adolescent substance use prevention programs do not work. *Evaluation Review* 17:529-55.
- Casswell, S., D. Mortimer, and C. Gilroy. 1982. The minimal effects and methodological problems in the evaluation of a harm reduction drug education programme in a high school setting. *Journal of Drug Education* 12:345-52.
- Center for Substance Abuse Prevention. 1993a. *Prevention strategies based on individual risk factors for alcohol and other drug abuse*. DHHS Publication No. ADM 1996-93. Washington, DC: U.S. Department of Health and Human Services.
- . 1993b. *Signs of effectiveness in preventing alcohol and other drug problems*. DHHS Publication No. SAM 93-2001. Washington, DC: U.S. Department of Health and Human Services.
- Clayton, R. R., A. M. Cattarello, and B. M. Johnstone. 1996. The effectiveness of Drug Abuse Resistance Education (Project D.A.R.E.): 5-year follow-up results. *Preventive Medicine* 25:307-18.
- Dunlap, E. and B. D. Johnson. 1992. The setting for the crack era: Macro forces, micro consequences (1960-1992). *Journal of Psychoactive Drug* 24:307-21.
- DuPont, P. J. and L. A. Jason. 1984. Assertiveness training in a preventive drug education program. *Journal of Drug Education* 14:369-78.
- Dusenbury, L. and M. Falco. 1995. Eleven components of effective drug abuse prevention curricula. *Journal of School Health* 65:420-5.
- Ellickson, P. L. 1995. Schools. In *Handbook on drug abuse prevention*, edited by R. H. Coombs and D. Ziedonis, 93-120. Boston: Allyn & Bacon.
- Ellickson, P. L. and R. M. Bell. 1990. Drug prevention in junior high: A multi-site longitudinal test. *Science* 247:1299-1305.
- . 1994. Ellickson and Bell respond. *American Journal of Public Health* 84:500-1.
- Ellickson, P. L., R. M. Bell, and K. McGuigan. 1993. Preventing adolescent drug use: Long-term results of a junior high program. *American Journal of Public Health* 83:856-61.
- Gerstein, D. R. and L. W. Green. 1993. *Preventing drug abuse: What do we know?* Washington, DC: National Academy Press.
- Goodstadt, M. S. 1980. Drug education—A turn on or a turn off? *Journal of Drug Education* 10:89-99.
- Gorman, D. M. 1994. Preventing adolescent drug use: The effectiveness of Project ALERT. *American Journal of Public Health* 84:500.
- . 1995. On the difference between statistical and practical significance in school-based drug abuse prevention. *Drugs: Education, Prevention and Policy* 2:275-83.
- . 1996. Etiological theories and the primary prevention of drug use. *Journal of Drug Issues* 26:505-20.
- . 1997. The failure of drug education. *Public Interest* 129:50-60.
- Gorman, D. M. and P. W. Speer. 1996. Preventing alcohol abuse and alcohol-related problems through community interventions: A review of evaluation studies. *Psychology and Health* 11:95-131.
- Hamid, A. 1991. From ganja to crack: Caribbean participation in the underground economy in Brooklyn, 1976-1986. Part 2. Establishment of the cocaine (and crack) economy. *International Journal of the Addictions* 26:729-38.
- Hansen, W. B. 1995. Project SMART. In *Encyclopedia of drugs and alcohol*. Vol. 2, edited by J. J. Jaffe, 873-76. New York: Simon & Schuster.

- Hansen, W. B., C. A. Anderson, B. R. Flay, J. W. Graham, and J. Sobel. 1988. Affective and social influence approaches to the prevention of multiple substance abuse among seventh grade students: Results from Project SMART. *Preventive Medicine* 17:135-54.
- Hansen, W. B. and J. W. Graham. 1991. Preventing alcohol, marijuana, and cigarette use among adolescents: Peer pressure resistance training versus establishing conservative norms. *Preventive Medicine* 20:414-30.
- Johnson, L. D. 1996. In *Marijuana use by youths continues to rise*, by C. S. Wren, *New York Times*, 20 February, A11.
- Kinder, B. N., N. E. Pape, and S. Walfish. 1980. Drug and alcohol education programs: A review of outcome studies. *International Journal of the Addictions* 15:1035-54.
- Mathias, R. 1994. School-based drug abuse prevention programs show long-lasting results. *NIDA News*, December, 8-9.
- McAlister, A. L. 1983. *Randomized study of tobacco and marijuana smoking prevention*. NIDA Contract Report, Center for Health Promotion, University of Texas, Houston.
- McAlister, A. L., C. Perry, J. Killen, L. A. Slinkard, and N. Maccoby. 1979. Pilot study of smoking, alcohol and drug prevention. *American Journal of Public Health* 70:719-21.
- Moskowitz, J. M., J. H. Malvin, G. A. Schaeffer, and E. Schaps. 1983. Evaluation of a junior high school primary prevention program. *Addictive Behaviors* 8:393-401.
- . 1984. An experimental evaluation of a drug education course. *Journal of Drug Education* 14:9-22.
- Moskowitz, J. M., E. Schaps, G. A. Schaeffer, and J. H. Malvin. 1984. Evaluation of a substance abuse prevention program for junior high school students. *International Journal of the Addictions* 19:419-30.
- Moynihan, D. P. 1966. What is "Community Action"? *Public Interest* 5:3-8.
- National Institute on Drug Abuse. 1997. *Preventing drug use among children and adolescents: A research-based guide*. Washington, DC: Author.
- Office of National Drug Control Policy. 1992. *National drug control strategy: A nation responds to drug use*. Washington, DC: The White House.
- . 1996. *National drug control strategy, 1996: Program, resources, and evaluation*. Washington, DC: The White House.
- Palinkas, L. A., C. J. Atkins, C. Miller, and D. Ferreira. 1996. Social skills training for drug prevention in high-risk female adolescents. *Preventive Medicine* 25:692-701.
- Pentz, M. A., J. H. Dwyer, D. P. MacKinnon, B. R. Flay, W. B. Hansen, E. Y. I. Wang, and C. A. Johnson. 1989. A multicommunity trial for primary prevention of adolescent drug abuse: Effects on drug use prevalence. *Journal of the American Medical Association* 261:3259-66.
- Randall, D. and W. R. Wong. 1976. Drug education to date: A review. *Journal of Drug Education* 6:1-21.
- Reinarman, C. and H. G. Levine. 1995. The crack attack: America's latest drug scare, 1986-1992. In *Images of issues: Typifying contemporary social problems*, edited by J. Best, 147-86. New York: Aldine De Gruyter.
- Riley details prospective cuts for safe and drug free schools. 1996. *Substance Abuse Funding News*, 12 February, 4-5.
- Rohrbach, L. A., C. D'Onofrio, T. E. Backer, and S. B. Montgomery. 1996. Diffusion of school-based substance abuse prevention programs. *American Behavioral Scientist* 39:919-34.
- Rosenbaum, D. P., R. L. Flewelling, S. L. Bailey, C. L. Ringwalt, and D. L. Wilkinson. 1994. Cops in the classroom: A longitudinal evaluation of Drug Abuse Resistance Education (D.A.R.E.). *Journal of Research in Crime and Delinquency* 31:3-31.

- Schaps, E., R. DiBartolo, J. Moskowitz, C. S. Palley, and S. Churgin. 1981. A review of 127 drug abuse prevention program evaluations. *Journal of Drug Issues* 11:17-43.
- Sowell, T. 1995. *The vision of the anointed: Self-congratulation as a basis for social policy*. New York: Basic Books.
- Substance Abuse and Mental Health Services Administration. 1996. *National household survey on drug abuse: Main findings 1994*. DHHS Publication No. SMA 96-3085. Washington, DC: U.S. Department of Health and Human Services.
- U.S. Senate Committee on Labor and Human Resources. 1990. *Drug and Alcohol Abuse Treatment and Prevention Improvement Act of 1990; Report to Accompany S. 2649*. Washington, DC: U.S. Government Printing Office.
- Van Biema, D. 1996. Just say life skills: A new antidrug program outstrips D.A.R.E. *Time*, 11 November, 70.
- Wren, C. S. 1996. Marijuana use by youths continues to rise. *The New York Times* 20 February, A11.
- Wren, C. S. 1997. New voice in drug debate seeks to lower the volume: Group calls for moderation and pragmatism. *The New York Times* 1 September, A10.

D. M. Gorman received his doctorate in sociology from the University of Essex in Great Britain. He is currently at the Center of Alcohol Studies and Department of Urban Studies and Community Health at Rutgers University. His primary research interest is in drug prevention policy.

SAMHSA SEEKS MEMBERS FOR REVIEW COMMITTEES

The Substance Abuse and Mental Health Services Administration's (SAMHSA) Office of Extramural Activities Review is soliciting curricula vitae (CVs) from persons who wish to be considered for membership on SAMHSA peer review committees. These committees review applications submitted for SAMHSA'S competitive grants and cooperative agreements, and proposals for contracts. Members of peer review committees provide expert advice on the scientific and technical quality of applications and proposals submitted to SAMHSA, and would be expected to prepare written critiques of their assigned applications and to attend a review meeting which may be held over a span of up to 5 days.

Candidates must have substantial expertise in the mental health and/or substance abuse prevention/treatment fields. In addition to candidates with broad based knowledge of the above fields, the Agency is especially interested in candidates who have substantial expertise in the following areas: (1) managed care models for treatment and prevention of substance abuse and mental disorders, (2) mental health and substance abuse issues of particular age groups (e.g., children/adolescents/seniors), ethnic/racial minority populations, and populations with special needs, (3) mental health issues related to disasters, (4) evaluation expertise in the mental health and/or substance abuse prevention/treatment fields. SAMHSA particularly wishes to ensure that women, ethnic/racial minorities, and persons with disabilities are adequately represented on its peer review committees.

Persons who submitted a CV previously need not resubmit these materials. Those names and CVs remain on file at SAMHSA. If selected for membership, you will be contacted by a Review Administrator regarding your interest and availability for a particular meeting in the future. You will also be provided with relevant information, including policies and procedures, meeting dates, and arrangements for travel to the Washington, DC, area, and the amount of per diem and honoraria provided.

CVs may be faxed to (301) 443-3437, or mailed to Dr. Jane Taylor, Director, Office of Extramural Activities Review, Room 17-89, 5600 Fishers Lane, Rockville, MD 20857. For additional information, call (301) 443-4783.

*A
Special
Issue
of
Evaluation
Review*

More Appropriate Evaluation Methods for Community-Level Health Interventions

Edited by John B. McKinlay

The field of evaluation research is undergoing rapid change. New sophisticated analytic approaches are emerging. Scientists with vision are urging a change in emphasis away from the study of individuals to the study of communities. The interest is not whether an individual quits smoking, or lowers their cholesterol level, but rather how many workplaces are designated as non-smoking, or how many restaurants offer heart-healthy menu items or whether air quality shows observable improvement, or is there a reduction in the overall rate of avoidable death?

The concept of process evaluation is relatively new, and receiving increased attention. Analyses of the methodology used in large scale community interventions may be as important as outcome measurement

The articles in this Special Issue of *Evaluation Review* consider different aspects of community-level evaluation research, and all demonstrate the emerging sophistication and promise of this field.

Contents: More Appropriate Evaluation Methods for Community-Level Health Interventions / Batch Sampling to Improve Power in a Community Trial: Experience From the Pawtucket Heart Health Program / A Strategy for Accurate Collection of Incremental Cost Data for Cost-Effectiveness Analyses in Field Trials: Pawtucket's Minimal Contact Cholesterol Education Intervention / Using Process Data to Explain Outcomes: An Illustration From the Child and Adolescent Trial for Cardiovascular Health (CATCH) / A Monte Carlo Study of Alternative Responses to Intraclass Correlation in Community Trials: Is It Ever Possible to Avoid Cornfield's Penalties? / Gauss or Bernoulli?

Evaluation Review

Volume 20, Number 3 / June 1996 / 120 pages

Single Issue: Individual \$49 / Institution \$16

Order Today!

Sage Customer Service: 805-499-9774 ■ **Sage FaxLine:** 805-499-0871

<http://www.sagepub.com>

SAGE PUBLICATIONS, INC.
2455 Teller Road
Thousand Oaks, CA 91320



SAGE PUBLICATIONS LTD
6 Bonhill Street
London EC2A 4PU, England

SAGE PUBLICATIONS INDIA PVT. LTD
M-32 Market, Greater Kailash I
New Delhi 110 048, India

A Special Issue of Evaluation & the Health Professions

MEDICAL OUTCOMES RESEARCH: IMPROVING MEASUREMENT IN MENTAL HEALTH

Edited by James S. Larson, *University of Arkansas*

The articles in this Special Issue make an important contribution to the literature on mental health outcomes research. It is an area of great importance to medical research, both from the viewpoint of curing illness and from the viewpoint of saving financial resources. It is hoped that the reader will find them useful and that they will stimulate further discussion in the future.

CONTENTS

The MOS 36-Item Short Form Health Survey:

A Conceptual Analysis

James S. Larson

The Proper Use of Self-Report Questions in Effective Measurement of Health Outcomes

Susan Schechter and Douglas Herrmann

Measuring Health Outcomes for Depression

Cathy Donald Sherbourne, Kenneth B. Wells, and Roland Sturm

Assessing the Effectiveness of Mental Health Care in Routine Clinical Practice:

Characteristics, Development, and Uses of Patient Outcomes Modules

G. Richard Smith, Jr., Kathryn M. Rost, Ellen P. Fischer,

M. Audrey Burnam, and Barbara J. Burns

Validation of the Panic Outcomes Module

*Jan Hollenberg, Kathryn M. Rost, Joylyn Humphrey, Richard R. Owen, Jr.,
and G. Richard Smith, Jr.*

An Instrument for Measurement of Outcomes of Care for Schizophrenia:

Issues in Development and Implementation

Brian J. Cuffel, Ellen P. Fischer, Richard R. Owen, Jr., and G. Richard Smith, Jr.

Improving Measurement in Mental Health

James S. Larson

Evaluation & the Health Professions

Volume 20, Number 1 / March 1997

Single Issue: Individual \$18 / Institution \$57

Order Today!

Sage Customer Service: 805-499-9774 • Sage FaxLine: 805-499-0871

Visit us at <http://www.sagepub.com>

SAGE PUBLICATIONS, INC.
2455 Teller Road
Thousand Oaks, CA 91320



SAGE PUBLICATIONS LTD
6 Bonhill Street
London EC2A 4PU, England

SAGE PUBLICATIONS INDIA PVT. LTD
M-32 Market, Greater Kailash I
New Delhi 110 048, India

*A leading journal of
quantitative research and methodology*

SMR/Sociological Methods & Research

Editor: Chris Winship, *Harvard University*

Devoted to the advancement of social and behavioral science methodology, **SMR/Sociological Methods & Research** emphasizes articles that clarify methodological problems, or that assist in organizing the known facts in a particular discipline or research area.

Empirically-based, **SMR/Sociological Methods & Research** strives to assess the validity, effectiveness and theoretical significance of the empirical research methods used in social science methodology. Though its articles are highly focused, the journal addresses a wide range of subject areas; sociology, psychology, political science, economics, anthropology, education research are just a few.

Features articles on such topics as:

- Simultaneous Latent Structure Analysis
- Principal Components Analysis
- Organizational Behavior Research
- Regression Analysis

Quarterly in August, November, February, May

Subscribe Today!			
Sage Customer Service: 805-499-9774 • Sage FaxLine: 805-499-0871			
	1 Year	2 Years	3 Years
Individual	\$73	\$146	\$219
Institutional	\$246	\$492	\$738

SAGE PUBLICATIONS, INC.
2455 Teller Road
Thousand Oaks, CA 91320

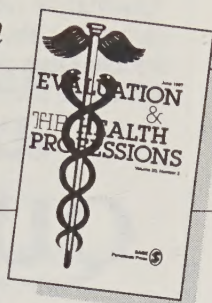


SAGE PUBLICATIONS LTD
6 Bonhill Street
London EC2A 4PU, England

SAGE PUBLICATIONS INDIA PVT. LTD
M-32 Market, Greater Kailash I
New Delhi 110 048, India

Your Key to the Latest Research

Evaluation & the Health Professions



Editor: R. Barker Bausell, *University of Maryland*

For nearly two decades, **Evaluation & the Health Professions** has provided a multidisciplinary forum for researchers, planners, and policymakers to examine the many applications of evaluation in health care programs and interventions.

Current and Comprehensive

Focusing exclusively on the development, implementation, and evaluation of health programs and interventions, **Evaluation & the Health Professions** offers immediate and potential applications to complex health care issues. Regular features include:

- ◆ Book reviews
- ◆ Original research
- ◆ Commentaries
- ◆ Reviews of evaluation efforts
- ◆ Methodological reports

Multidisciplinary Perspective

With **Evaluation & the Health Professions**, distinguished scholars and practitioners from a variety of health fields present a broad, multidisciplinary perspective through original research designed to define health care programs and impact practice and policy. Recent articles have discussed:

- ◆ Curriculum evaluation and development
- ◆ Estimating cancer screening indicators
- ◆ Evaluating public health service programs
- ◆ The hospital culture scale
- ◆ Legal vulnerability
- ◆ Measuring perinatal drug treatment success
- ◆ No-fault medical liability

Subscribe Today!

Sage Customer Service: 805-499-9774 • **Sage FaxLine:** 805-499-0871

Visit us at <http://www.sagepub.com>

	<i>1 Year</i>	<i>2 Years</i>	<i>3 Years</i>
Individual	\$64	\$128	\$192
Institutional	\$219	\$438	\$657

SAGE PUBLICATIONS, INC.
2455 Teller Road
Thousand Oaks, CA 91320



SAGE PUBLICATIONS LTD
6 Bonhill Street
London EC2A 4PU, England

SAGE PUBLICATIONS INDIA PVT. LTD
M-32 Market, Greater Kailash I
New Delhi 110 048, India

"Qualitative researchers of all persuasions need a place to talk to each other, thoughtfully and non-polemically, about their craft and the ideas that underlie it. **Qualitative Inquiry** is that place – and we will all be the richer for it."

—Matt Miles, *Senior Research Associate, Center for Policy Research, New York, Co-Author of Qualitative Data Analysis*

QUALITATIVE INQUIRY

Editors: Norman K. Denzin, *University of Illinois,
Urbana-Champaign*

Yvonna S. Lincoln, *Texas A&M University*

Taking the lead, Sage Periodicals Press is proud to be the publisher of **Qualitative Inquiry**, the first interdisciplinary journal to provide a forum for qualitative methodology and related issues in the human sciences.

Qualitative Inquiry concentrates on qualitative research that transcends disciplinary, racial, ethnic, gender, national, and paradigmatic boundaries. **QI** includes research that experiments with nontraditional forms, content, and modes of representation. Periodic Special Issues will provide integrated, focused treatment of topics of contemporary concern.

Under the leadership of editors Norman K. Denzin and Yvonna S. Lincoln, **Qualitative Inquiry** publishes articles addressing research issues, paradigmatic questions, methodological strategies and techniques, syntheses of qualitative work in broad substantive areas, and practical applications of qualitative methodologies. Results of specific research studies will be included only when they are paramount to the methodologies employed.

SUBSCRIBE TODAY!

Sage Customer Service: 805-499-9774 • Sage FaxLine: 805-499-0871

Visit us at <http://www.sagepub.com>

SAGE PUBLICATIONS, INC.
2455 Teller Road
Thousand Oaks, CA 91320



SAGE PUBLICATIONS LTD
6 Bonhill Street
London EC2A 4PU, England

SAGE PUBLICATIONS INDIA PVT. LTD
M-32 Market, Greater Kailash I
New Delhi 110 048, India

Now bimonthly!

EDUCATIONAL and PSYCHOLOGICAL MEASUREMENT

Now a bimonthly journal devoted to the development and application of measures of individual differences.

Editor: Bruce Thompson, Texas A & M University

For more than five decades, social scientists have depended on **Educational and Psychological Measurement** for the latest systematic investigations concerned with the validity of educational and psychological measures. With six issues per volume, **Educational and Psychological Measurement**, every volume will bring you even more of the latest research on the development application of measures of individual differences.

Whether your studies concern testing and measurement methods in education, psychology, industry, or government, **Educational and Psychological Measurement** will now bring you more of the in-depth scholarly research you need to stay on top of this dynamic and vital field.

The special features of **Educational and Psychological Measurement** are:

- ◆ Refereed, in-depth research pieces
- ◆ Validity Studies, providing early publication of short studies dealing with validity of tests and scales
- ◆ Book Reviews featuring important new books of interest to evaluation and methodology researchers
- ◆ Computer Studies dedicated to the timely publication of computer programs relevant to the fields of education and psychology

Order Today!

Sage Customer Service: 805-499-9774 ■ Sage FaxLine: 805-499-0871
Visit us at <http://www.sagepub.com>

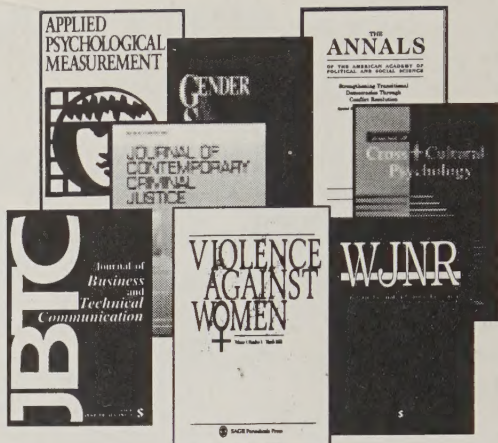
SAGE PUBLICATIONS, INC.
2455 Teller Road
Thousand Oaks, CA 91320



SAGE PUBLICATIONS LTD
6 Bonhill Street
London EC2A 4PU, England

SAGE PUBLICATIONS INDIA PVT. LTD
M-32 Market, Greater Kailash I
New Delhi 110 048, India

4 ways to order and share the best scholarship in your field



■ Back Issues

Make sure you didn't miss any important articles that you need for your own practice, research or classroom. Back issues are available at 20% discount. To order, call Sage Customer Service at (805) 499-9774 or email: order@sagepub.com.

■ Reprints

Customize your instructional materials to fit your classroom and symposia needs with powerful, new ideas from leading scholars in the field. Reprints make cost-effective additions to your course materials and conference packets. They are convenient and easy to obtain. To order multiple copies of a recent article, call Sage Reprint Services at (805) 499-0721, Ext 7535 or email: reprint@sagepub.com.

■ Special Issues

Occasionally an entire issue is devoted to a single relevant topic, viewing it at length and in depth. The pertinent information in these Special Issues is something you will refer to again and again. Sage offers discounts of up to 60% on multiple copies for classroom adoption. For more information on multiple-copy prices call (805) 499-0721, ext. 7528 or email: adopt_jnl@sagepub.com.

■ Sample Copies

Sample copies (up to two) may be obtained on request directly from Sage Publications, by calling Sage Customer Service at (805) 499-9774 or e-mail: jsamples@sagepub.com.



SAGE PUBLICATIONS, INC.

2455 Teller Road, Thousand Oaks, CA 91320

Tel: (805) 499-0721 ■ Fax: (805) 499-0871

Visit us at <http://www.sagepub.com>